Montana Science Assessment

2022–2023

Volume 1:
Annual Technical Report

Office of Public Instruction
Elsie Arntzen, State Superintendent

TABLE OF CONTENTS

LIST OF TABLES

List of Figures

List of Appendices

1. INTRODUCTION

The Montana Science Assessment (MSA) is a science assessment for grades 5 and 8. The *Montana Science Assessment Technical Report 2022-2023* is provided to document and make transparent all methods used in item development, test construction, psychometrics, standard setting, test administration, and score reporting, including summaries of student results and evidence and support for the intended uses and interpretations of the test scores. The technical reports are reported as six separate, self-contained volumes, as described in the following list:

1) **Annual Technical Report.** This volume is updated each year and provides a global overview of the tests administered to students annually.

2) **Test Development.** This volume summarizes the procedures used to construct test forms and provides summaries of the item bank and development process.

3) **Setting Performance Standards.** This volume documents the methods and results of the MSA standard-setting process.

4) **Evidence of Reliability and Validity.** This volume provides technical summaries of the test quality and special studies conducted to support the intended uses and interpretations of the test scores.

5) **Test Administration.** This volume describes the security protocols, accessibility features (including accommodations), methods used, and system characteristics developed to administer tests.

6) **Score Interpretation Guide.** This volume describes the score types reported and details the appropriate inferences that can be drawn from each score reported.

The Montana Office of Public Instruction (OPI) communicates the quality of the MSA by making these technical reports accessible to the public on the state's website.

1.1 BACKGROUND AND HISTORICAL CONTEXT OF TESTS

In 2016, Montana adopted a new set of K–12 science standards as the Montana Science Content Standards. The new standards use the three-dimensional framework in *A Framework for K–12 Science Education* (National Research Council, 2012) as the foundation, and closely resemble the Next Generation Science Standards (NGSS). The three-dimensional science standards reflect the latest research and advances in modern science education. They describe specific performances that demonstrate what students know and can do. Each standard incorporates three dimensions: a science or engineering practice, a disciplinary core idea, and a crosscutting concept. The OPI and its assessment vendor, Cambium Assessment, Inc. (CAI), developed and administered a new online assessment to measure these new standards. Field-tested in 2020–2021 and administered operationally for the first time in 2021–2022, the MSA measures the science knowledge and skills of Montana students in grades 5 and 8.

The OPI provides an overview of the science assessment at:
https://opi.mt.gov/Leadership/Assessment-Accountability/MontCAS/Required-

Assessments/Montana-Science-Assessment-FAQ. Information about the Montana Science Content Standards is available at: https://opi.mt.gov/Educators/Teaching-Learning/K-12-Content-Standards/Science-Standards.

1.2 Purpose and Intended Uses of the Montana Science Assessment

The MSA is a criterion-referenced test established using principles of evidence-centered design to yield overall and discipline-level test scores at the student level and other levels of aggregation that reflect student achievement of the MSA. Volume 2 of this technical report details test development. The three-dimensional science standards (i.e., the Montana Science Content Standards) establish a set of knowledge and skills that all students need to be prepared for a wide range of high-quality post-secondary opportunities, including higher education and entering the workplace.

The three-dimensional science standards reflect the latest research and advances in modern science education and differ from previous science standards in multiple ways. First, rather than describing general knowledge and skills that students should know and be able to do, they describe specific performances that demonstrate what students know and can do. The Montana Science Content Standards refer to such performed knowledge and skills as *performance expectations* (PEs). Second, the Montana Science Content Standards are intentionally multi-dimensional. Each performance expectation (PE) incorporates all three dimensions from *A Framework for K–12 Science Education* (National Research Council, 2012)—a science or engineering practice, a disciplinary core idea, and a crosscutting concept. Another unique feature of the Montana Science Content Standards is the assumption that students should learn all science disciplines rather than a select few, as is traditionally done in many high schools, where students may elect, for example, to take biology and chemistry but not physics or astronomy.

The MSA supports instruction and student learning by providing valuable feedback to educators and parents, which can be used to form instructional strategies to remediate or enrich instruction. An array of reporting metrics is provided to evaluate performance at the student and aggregate levels and to monitor improvement at the student and group levels over time. Volume 6 of this technical report details score reporting.

The MSA draws items from an item bank that consists of Independent College and Career Readiness (ICCR) items and items owned by several other states and one U.S. territory that abide by a Memorandum of Understanding (MOU) to share content, leadership, and new ideas and methods. Full members of the MOU in 2023 were Connecticut, Hawaii, Idaho, Montana, Oregon, Rhode Island, Utah, West Virginia, and Wyoming. CAI had a supporting and coordinating role. New Hampshire, North Dakota, South Dakota, and U.S. Virgin Islands observed and participated in some activities. CAI and the OPI worked together to ensure that the items in the tests constructed for all grades uniquely measured students' mastery of the three-dimensional Montana Science Content Standards.

Table 1 outlines the required uses and citations for the MSA based on the Administrative Rules of Montana (ARM) and the federal *Every Student Succeeds Act* (ESSA) plan. The MSA fulfills all the requirements described in Table 1.

Table 1. Required Uses and Citations for the Montana Science Assessment

Required Use	Required Use Citation
Indicator of academic achievement and progress	ESSA section 1111(b)(2)(B)(ii)
Test administration frequency and grade levels	ESSA section 1111(b)(2)(B)(v)(II) ARM 10.56.102(3)
Disaggregation of test scores	ESSA section 1111(b)(2)(B)(xi)
Publication of test scores	ESSA section 1111(b)(2)(B)(x) ARM 10.56.102(6)(c)
Requirement of the alignment of test to academic content standards	ARM 10.56.102(3)(b)

1.3 Participants in the Development and Analysis of the Montana Science Assessment

The OPI manages the MSA programs with the assistance of several participants, including Montana educators, a Technical Advisory Committee (TAC), and vendors. The OPI fulfills the diverse requirements of implementing Montana's statewide assessments while adhering to the guidelines established in the *Standards for Educational and Psychological Testing* (American Educational Research Association [AERA], American Psychological Association [APA], & National Council on Measurement in Education [NCME], 2014). To comply with the *Standards*, scale development, scoring, linking, and evaluation of differential item functioning are addressed in the current volume; item development, test design, and test blueprints are documented in Volume 2, Test Development; development of cut scores is summarized in Volume 3, Setting Performance Standards; evidence for validity and reliability/precision was collected and is reported in Volume 4, Evidence of Reliability and Validity; information on testing windows, test options, accommodations, training of test coordinators and administrators, and test security are provided in Volume 5, Test Administration; supporting documentation for tests, score uses and interpretation are included in Volume 6, Score Interpretation.

1.3.1 Montana Office of Public Instruction

The Assessment Unit manages test development, administration, scoring, and results reporting for the statewide comprehensive assessment programs, including coordinating with other OPI offices, Montana public schools, and vendors.

1.3.2 Montana Educators

Montana educators participate in most aspects of the conceptualization and development of the MSA. Educators participate in clarifying how the standards are assessed, designing tests, and reviewing test questions and passages.

1.3.3 Technical Advisory Committee

The OPI convenes an advisory committee panel twice each year to discuss psychometrics, test development, and administrative and policy issues relevant to the current and future Montana assessments. This committee is comprised of several nationally recognized assessment experts and highly experienced practitioners from several Montana school districts.

1.3.4 Cambium Assessment, Inc.

CAI (formerly the American Institutes for Research [AIR]) is the vendor that was selected through the state-mandated competitive procurement process. CAI is responsible for developing test content, building test forms, conducting psychometric analyses, administering and scoring test forms, and reporting test results for the MSA described in this report. Additionally, CAI is responsible for developing and maintaining the ICCR item bank.

1.3.5 Caveon Test Security

Caveon Test Security monitored web pages and social media during the spring 2023 test administration to ensure that secure testing materials such as items and prompts were not leaked in the Spring 2023 MSA administration. Details of Caveon Test Security are described in Appendix 1-A, Caveon Test Security Overview.

1.4 Available Test Formats and Special Versions

The MSA is administered online using an adaptive test design; Section 3.3 of this volume details the adaptive test design. Science items are centered on a scientific phenomenon. They can consist of shorter (stand-alone) items or items with several parts (item clusters) requiring the student to interact with them in various ways. The science test was administered as an independent field test in spring 2021 and went operational in spring 2022. Starting in 2022 and going forward, additional items were field-tested to build out the item bank.

Students unable to participate in the online test administration have the option to use print-on-demand—a feature that provides the same items administered to students online in a paper-pencil format. Spanish versions of the MSA (developed to meet the same content standards as the English versions) are available for all tested grades. Students participating in the computer-based MSA can use standard online testing features in the Test Delivery System (TDS), including a selection of font colors and sizes and the ability to zoom in and out or highlight text. In addition to the resources available to all students, options are available to accommodate students with an Individualized Education Program (IEP) or Section 504 Plan. These include braille, American Sign Language (ASL), closed captioning, and large print. Students with disabilities have the option to take the MSA with or without accommodations or to take an alternate assessment. For additional information about testing features and accommodations, refer to Volume 5, Test Administration, of this technical report.

1.5 Student Participation

The MSA is administered in the spring. All students (including retained students) enrolled in grades 5 and 8 at public schools in Montana are required to participate in the MSA. Table 2 shows the number of students who were tested (Number Tested) and the number of students whose scores were included for analyses in this technical report (Number Reported). In Spring 2023, participation rates were XX% and YY% for G5 and G8.

Table 3 shows the demographic characteristics of the student population, by counts and percentages, in the spring administration of the 2022–2023 MSA. The subgroups reported are gender, ethnicity, English learners, special education students, and economically disadvantaged students.

Table 2. Number of Students Participating in the Montana Science Assessment, Spring 2023

Grade	Number Tested	Number Reported
5	11,070	11,068
8	11,164	11,159

Table 3. Distribution of Demographic Characteristics of Student Population

Group	Grade 5		Grade 8	
	N	*%*	*N*	*%*
All Students	11,068	100.00	11,159	100.00
Female	5,406	48.84	5,384	48.25
Male	5,662	51.16	5,775	51.75
African American	88	0.80	57	0.51
American Indian/Native Alaskan	1,089	9.84	1,096	9.82
Asian	80	0.72	77	0.69
Hispanic	609	5.50	660	5.91
Multi-Racial	536	4.84	532	4.77
Pacific Islander	16	0.14	25	0.22
White	8,650	78.15	8,712	78.07
English Learners	374	3.38	458	4.10
Special Education	1,452	13.12	1,292	11.58
Economically Disadvantaged	4,856	43.87	4,070	36.47

2. Operational Practices and Procedures

This section outlines key elements of the operational administration, including testing window, test administrators, online testing environment, and simulations. Accessibility supports including universal tools, designated supports, and accommodations are also discussed, followed by the number of test sessions with allowed designated supports and accommodations for each test.

2.1 Testing Window

Table 4 shows the testing windows for the 2022–2023 Montana Science Assessment (MSA).

Table 4. Montana Science Assessment Testing Windows

Grade	Start Date	End Date
5, 8	3/6/2023	5/26/2023

2.2 Test Administrators

Key personnel involved with test administration for the Montana Office of Public Instruction (OPI) included district test coordinators (DCs), school test coordinators (SCs), and test administrators (TAs) who proctored the test. Test Administration Manuals (TAM; refer to Appendix 5-A in Volume 5, Test Administration) were provided so that personnel involved with the statewide assessment administrations could maintain both standardized administration conditions and test security.

2.3 Testing Environment

The Cambium Assessment, Inc. (CAI) Secure Browser was required to access the online assessment. The online browser provided a secure environment for student testing by disabling the hot keys and copy and screen capture capabilities, and preventing access to the desktop (Internet, email, and other files or programs installed on school machines). During the online assessment, students could pause a test, review previously answered questions, and modify their responses if the test had not been paused for more than 20 minutes. Students do not have a required time limit for each test session, but for test administration planning purposes, schools were given approximate time estimates for how long each test may take. For additional information about the test administration, refer to Volume 5, Test Administration, of this technical report.

2.4 Simulations

CAI employs a simulation approach to all MSA tests before the operational testing window begins. For adaptive tests, simulations were conducted to configure the item selection algorithm settings, evaluate whether individual tests adhered to the test blueprint and correlated highly with student ability, to monitor item exposure rates, and to verify the scores produced by CAI's scoring engine. Simulations were also conducted on fixed-form tests to quality check the scores. The simulation approaches and results are discussed in Volume 2, Test Development, of this technical report.

2.5 Universal Tools, Designated Supports, and Accommodations

Accessibility supports are available to students when needed to remove barriers during testing while maintaining the constructs that are measured by the MSA. The accessibility supports discussed in this technical report include embedded (digitally provided) and non-embedded (non-digitally or locally provided) universal features available to all students as they access instructional or assessment content; designated supports available to those students for whom the need has been identified by an informed educator or team of educators; and accommodations generally available for students for whom there is documentation on an Individualized Education Program (IEP) or Section 504 Plan. For English learners (ELs), Spanish language versions of the MSA were available.

All educators making decisions about designated supports were trained on the process and understand the range of designated supports available.

Accommodations are changes in procedures or materials that ensure equitable access to instructional and assessment content and generate valid assessment results for students who need them. Embedded accommodations (e.g., text-to-speech [TTS]) are provided digitally through instructional or assessment technology, and non-embedded designated features (e.g., scribe) are non-digital. State-approved accommodations do not compromise the learning expectations, constructs, or grade-level standards. These accommodations help students with a documented need generate valid assessment outcomes that fully demonstrate what they know and are able to do. From the psychometric point of view, the purpose of providing accommodations is to "increase the validity of inferences about students with special needs by offsetting specific disability-related, construct-irrelevant impediments to performance" (Koretz & Hamilton, 2006, p. 562).

Montana TAs and SCs were responsible for ensuring that arrangements for accommodations were made before the test administration dates. The available accommodation options for eligible students include braille, American Sign Language (ASL), closed captioning, streamline, abacus, assistive technology (e.g., adaptive keyboards, touch screens, switches), calculation device, print-on-demand, multiplication table, and scribe.

Additional information about universal features, designated supports, and accommodations can be found in Volume 5, Test Administration, of this technical report.

Table 5 and Table 6 list the number of test sessions in which students used each of the designated supports or accommodations during the spring 2023 test administration.

Table 5. Number of Testing Sessions with Allowed Designated Supports

Designated Supports	**Grade**			
	5		**8**	
	N	**%**	**N**	**%**
Embedded				
Color Choices	27	0.24%	15	0.13%
Language (Spanish)	8	0.07%	15	0.13%
Masking	78	0.70%	47	0.42%
Mouse Pointer	21	0.19%	2	0.02%
Permissive Mode	7	0.06%	16	0.14%
Streamlined Mode	60	0.54%	30	0.27%
Text-to-Speech	970	8.76%	712	6.38%

Designated Supports	**Grade**			
	5		**8**	
	N	**%**	**N**	**%**
Zoom	13	0.12%	11	0.10%
Non-Embedded				
Amplification	-	-	-	-
Bilingual Dictionary	-	-	-	-
Color Contrast	21	0.19%	2	0.02%
Color Overlay	12	0.11%	6	0.05%
Magnification	-	-	-	-
Medical Supports	-	-	-	-
Noise Buffers	79	0.71%	47	0.42%
Read Aloud	-	-	-	-
Read Aloud Spanish	4	0.04%	4	0.04%
Scribe Items (non-writing)	93	0.84%	47	0.42%
Separate Setting	854	7.72%	672	6.02%
Simplified Test Directions	-	-	-	-
Translated Test Directions	-	-	-	-

Table 6. Number of Testing Sessions with Allowed Accommodations

Accommodations	**Grade**			
	5		**8**	
	N	**%**	**N**	**%**
Embedded				
Braille Type	-	-	-	-
Language (Braille)	-	-	-	-
Print on Request	25	0.23%	10	0.09%
Non-Embedded				
Alternate Response Options	3	0.03%	1	0.01%
Sign Language - Test Items	-	-	1	0.01%
Specialized Calculator	33	0.30%	148	1.33%
Speech-to-Text	108	0.98%	103	0.92%
Timing or Scheduling	11	0.10%	14	0.13%
Word Prediction	69	0.62%	66	0.59%

3. ITEM BANK AND TEST DESIGN

3.1 SHARED SCIENCE ASSESSMENT ITEM BANK

Cambium Assessment, Inc. (CAI) works with a group of states and one U.S. territory to develop science assessments to assess the Next Generation Science Standards (NGSS) and other standards influenced by the same science framework. Many of these states have signed a Memorandum of Understanding (MOU) to share item specifications and items. CAI coordinates this group of states and holds contracts to develop and deliver the items for most of them.

CAI also built the Independent College and Career Readiness (ICCR) science item pool in partnership with these states and one U.S. territory. These CAI-owned items make up a substantial part of the item bank and are shared with partner states and one U.S. territory. Montana signed the MOU, and therefore, the item pool available for the Montana Science Assessment includes items from the following three sources:

1. Items owned by Montana
2. Items shared by other states/territory within the MOU collaboration
3. Items shared from the ICCR item bank

In 2023, the Shared Science Assessment Item Bank was used for operational tests in 12 states and one U.S. territory, including Montana. The goals, uses, and claims that the Shared Science Assessment Item Bank and resulting tests are designed to support were identified in a collaborative meeting on August 22–23, 2016, to facilitate the transition from a framework for three-dimensional science standards, specifically the NGSS, to statewide summative assessments for science. CAI invited content and assessment leaders from 10 states and four nationally recognized experts who helped co-author the NGSS. Two nationally recognized psychometricians also participated.

In 2017, cognitive lab studies were conducted to evaluate and refine the process of developing item clusters aligned to the three-dimensional science standards. The results of the cognitive lab studies confirmed the feasibility of the approach (refer to Volume 4, Appendix 4-D, Science Clusters Cognitive Lab Report, of this technical report).

A second set of cognitive lab studies was conducted in 2018 and 2019 to determine whether students using braille could understand the task demands of selected accommodated three-dimensional science-aligned item clusters. They also evaluated whether these students could navigate the interactive features of these item clusters in a manner that allowed them to fully display their knowledge and skills relative to the constructs of interest. In general, both the students who relied entirely on braille and/or Job Access With Speech (JAWS) and those who had some vision and were able to read the screen with magnification were able to find the information they needed to respond to the questions, navigate the various response formats, and finish within a reasonable amount of time (refer to Volume 4, Appendix 4-E, Braille Cognitive Lab Report, of this technical report).

In 2018, CAI field tested 394 item clusters and stand-alone items in elementary and middle school, of which 338 (including items from all sources) were accepted and made available as operational items in 2019. The numbers of item clusters and stand-alone items that were field tested were:

244 in 2019, 373 in 2021, and 360 in 2022, while the numbers of items that were accepted and made available for future operational use were 185 in 2019, 317 in 2021, and 313 in 2022. In 2023, 259 item clusters and stand-alone items in elementary and middle school were field tested, of which 224 (17 of which are Montana-owned) were accepted and made available for operational use in the future years. All these items follow the same specifications, test development processes, and review processes, all of which involve the following procedures:

- CAI staff and participating states collaborated to develop item specifications, which are documents designed to guide item writers as they craft test questions and stakeholders as they review those items. The item specifications were generally accompanied by sample items meeting those specifications. All specifications and sample items were reviewed by state content experts and committees of educators in at least one state.
- The specifications helped test developers create item clusters and stand-alone items that covered a range of difficulty, furthering the goal of measuring the full range of performance found in the population, but remaining at grade level. All item writers were trained in the principles of universal design, the appropriate use of item interactions, and the science item specifications.
- Items were reviewed by science experts in at least one state.
- Every item was reviewed by a content advisory committee (comprised of state educators) in at least one state or in a cross-state educator review process.
- Every item was reviewed by a committee of educators charged with evaluating language accessibility, bias, and sensitivity in at least one state or a cross-state educator review.
- Every item was field tested, all scoring protocols (i.e., rubrics) were validated using the field-test data, and items with questionable data were reviewed again by committees of educators.

A detailed description of the Shared Science Assessment Item Bank development process is included in Volume 2, Test Development, of this technical report.

3.2 FIELD-TESTING

All items that were part of the operational pool of the Shared Science Assessment Item Bank were field tested in prior years, which is documented in Appendix 1-B, Shared Science Assessment Item Bank: Field-Testing. Field-testing for the current administration is described in this section.

3.2.1 2023 Field Tests

In 2023, items were field tested in 12 states and one U.S. territory. In all 12 states and one U.S. territory (Connecticut, Hawaii, Idaho, Montana, New Hampshire, North Dakota, Oregon, Rhode Island, South Dakota, U.S. Virgin Islands, Utah, West Virginia, and Wyoming), field-test items were administered as unscored items embedded among operational items. In total, 119 item clusters and 143 stand-alone items were administered as field-test items in the elementary and middle school grade bands. Table 7 presents the number of field-test item clusters and stand-alone items administered in each grade band for each state. The numbers in parentheses in the column representing Montana show the number of field-test items owned by Montana.

Table 7. Number of Field-Test Items Administered, Spring 2023

Grade Band and Item Type	CT	HI	ID	MT	NH	ND	OR	RI	SD	USVI	UT	WV	WY	Total*
Elementary School	**41**	**19**	**20**	**12 (12)**	**11**	**12**	**28**	**35**	**10**	**1**	**32**	**25**	**7**	**126**
Cluster	16	7	12	4 (4)	7	4	12	12	4	1	32	7	3	60
Stand-Alone	25	12	8	8 (8)	4	8	16	23	6			18	4	66
Middle School	**36**	**24**	**21**	**9 (9)**	**11**	**11**	**41**	**29**	**7**	**1**	**49**	**28**	**7**	**136**
Cluster	6	8	5	5 (5)	7	4	10	9	5	1	49	8	3	59
Stand-Alone	30	16	16	4 (4)	4	7	31	20	2			20	4	77
Total	**77**	**43**	**41**	**21 (21)**	**22**	**23**	**69**	**64**	**17**	**2**	**81**	**53**	**14**	**262**

*The total count excludes 82 Oregon and South Dakota legacy field-tested items in elementary school and middle school, 6 HI enemy pilot study field-tested items in elementary school and middle school, but includes several field-tested items moved to comprehensive interim pool after rubric validation.

Two of the states (New Hampshire and Rhode Island) opted for a test in which operational items were grouped by science discipline. For these two states, the field-test items were presented together in a fourth group of items. The sequence of the four sets of items (corresponding to the three disciplines and a set of field-test items) was randomized across students. Ten other states and one U.S. territory (Connecticut, Hawaii, Idaho, Montana, North Dakota, Oregon, South Dakota, Utah, U.S. Virgin Islands, West Virginia, and Wyoming) opted for a test design in which the items were not grouped by discipline. In these 10 states and one U.S. territory, field-test items were administered at random positions throughout the test. A student received either a field-test item cluster or a set of four field-test stand-alone items. The test design for the MSA is discussed in Section 3, Item Bank and Test Design of this volume.

A minimum sample size of 1,500 students per field-test item was targeted for any given state or territory. Most items were administered in two states or one territory. Approximately 85.5% of the items met or exceeded the target sample size of 1,500 in at least one state, and all items had a sample size of at least 1,350 (90% of the target) in at least one state.

Table 8 and Table 9 present the number of item clusters and stand-alone items that were shared between the field-test pools of any two states or territory. The numbers below the shaded cells represent the number of common field-test items between any two states, and the numbers above the shaded cells represent the number of common field-test items that survived rubric validation and were included in the calibration. In each of the shaded cells, the number outside the parentheses represents the number of unique field-test items administered only in the given state or territory, and the number in the parentheses represents the number of unique and/or common items that were calibrated with only the data from that state. Table 8 presents the results for elementary schools, Table 9 presents the results for middle schools. The numbers of field-test items administered are slightly different from the numbers of field-test items at calibration because some items were rejected during rubric validation.

Table 8. Number of Common Elementary School Field-Test Items Administered and Calibrated, Spring 2023

	State	CT	HI	ID	MT	NH	ND	OR	RI	SD	USVI	UT	WV	WY
Item Clusters	CT	0 (0)	0	2	2	2	0	5	4	0	0	1	0	0
	HI	0	0 (0)	1	0	0	0	0	0	0	0	6	0	0
	ID	2	1	0 (0)	0	0	0	0	2	0	0	6	0	0
	MT	2	0	0	0 (0)	0	0	0	0	0	0	2	0	0
	NH	2	0	0	0	0 (0)	0	0	0	0	0	2	3	0
	ND	0	0	0	0	0	0 (0)	3	0	0	0	1	0	0
	OR	5	0	0	0	0	3	0 (0)	0	0	0	4	0	0
	RI	4	0	2	0	0	0	0	0 (0)	0	0	4	2	0
	SD	0	0	0	0	0	0	0	0	0 (0)	1	2	2	0
	USVI	0	0	0	0	0	0	0	0	1	0 (0)	1	0	0
	UT	1	6	7	2	2	1	4	4	2	1	0 (0)	0	3
	WV	0	0	0	0	3	0	0	2	2	0	0	0 (0)	0
	WY	0	0	0	0	0	0	0	0	0	0	3	0	0 (0)
Stand-Alone Items	CT	0 (0)	0	6	3	0	0	2	9	0	0	0	3	0
	HI	0	0 (0)	0	0	0	0	8	0	0	0	0	0	4
	ID	6	0	0 (0)	0	0	0	0	2	0	0	0	0	0

	State	CT	HI	ID	MT	NH	ND	OR	RI	SD	USVI	UT	WV	WY
	MT	4	0	0	0 (0)	0	0	0	4	0	0	0	0	0
	NH	0	0	0	0	0 (0)	0	0	0	0	0	0	4	0
	ND	0	0	0	0	0	0 (0)	6	2	0	0	0	0	0
	OR	2	8	0	0	0	6	0 (0)	0	0	0	0	0	0
	RI	10	0	2	4	0	2	0	0 (0)	0	0	0	5	0
	SD	0	0	0	0	0	0	0	0	0 (0)	0	0	6	0
	USVI	0	0	0	0	0	0	0	0	0	0 (0)	0	0	0
	UT	0	0	0	0	0	0	0	0	0	0	0 (0)	0	0
	WV	3	0	0	0	4	0	0	5	6	0	0	0 (0)	0
	WY	0	4	0	0	0	0	0	0	0	0	0	0	0 (0)
Total	CT	0 (0)	0	8	5	2	0	7	13	0	0	1	3	0
	HI	0	0 (0)	1	0	0	0	8	0	0	0	6	0	4
	ID	8	1	0 (0)	0	0	0	0	4	0	0	6	0	0
	MT	6	0	0	0 (0)	0	0	0	4	0	0	2	0	0
	NH	2	0	0	0	0 (0)	0	0	0	0	0	2	7	0
	ND	0	0	0	0	0	0 (0)	9	2	0	0	1	0	0

	State	CT	HI	ID	MT	NH	ND	OR	RI	SD	USVI	UT	WV	WY
	OR	7	8	0	0	0	9	0 (0)	0	0	0	4	0	0
	RI	14	0	4	4	0	2	0	0 (0)	0	0	4	7	0
	SD	0	0	0	0	0	0	0	0	0 (0)	1	2	8	0
	USVI	0	0	0	0	0	0	0	0	1	0 (0)	1	0	0
	UT	1	6	7	2	2	1	4	4	2	1	0 (0)	0	3
	WV	3	0	0	0	7	0	0	7	8	0	0	0 (0)	0
	WY	0	4	0	0	0	0	0	0	0	0	3	0	0 (0)

Table 9. Number of Common Middle School Field-Test Items Administered and Calibrated, Spring 2023

	State	CT	HI	ID	MT	NH	ND	OR	RI	SD	USVI	UT	WV	WY
Item Clusters	CT	0 (0)	0	0	0	0	0	0	1	0	1	4	1	0
	HI	0	0 (0)	0	0	0	0	0	0	0	0	8	0	0
	ID	0	0	0 (0)	0	0	0	0	1	0	0	3	2	0
	MT	0	0	0	0 (0)	0	0	0	0	0	0	5	0	0
	NH	0	0	0	0	0 (0)	0	0	0	0	0	7	0	0
	ND	0	0	0	0	0	0 (0)	0	1	0	0	3	0	0
	OR	0	0	0	0	0	0	0 (0)	1	0	0	8	1	0
	RI	1	0	1	0	0	1	1	0 (0)	2	0	2	2	0
	SD	0	0	0	0	0	0	0	2	0 (0)	0	3	0	0
	USVI	1	0	0	0	0	0	0	0	0	0 (0)	1	0	0
	UT	4	8	3	5	7	3	8	2	3	1	0 (0)	3	2
	WV	1	0	2	0	0	0	1	2	0	0	3	0 (0)	0
	WY	0	0	0	0	0	0	0	0	0	0	3	0	0 (0)
Stand-Alone Items	CT	0 (0)	4	10	0	4	0	7	0	0	0	0	4	0
	HI	4	0 (0)	5	0	0	0	3	0	0	0	0	4	0
	ID	10	5	0 (0)	0	0	0	0	1	0	0	0	0	0

	State	CT	HI	ID	MT	NH	ND	OR	RI	SD	USVI	UT	WV	WY
	MT	0	0	0	0 (0)	0	0	4	0	0	0	0	0	0
	NH	4	0	0	0	0 (0)	0	0	0	0	0	0	0	0
	ND	0	0	0	0	0	0 (0)	0	7	0	0	0	0	0
	OR	8	3	0	4	0	0	0 (0)	6	0	0	0	7	3
	RI	0	0	1	0	0	7	6	0 (0)	2	0	0	4	0
	SD	0	0	0	0	0	0	0	2	0 (0)	0	0	0	0
	USVI	0	0	0	0	0	0	0	0	0	0 (0)	0	0	0
	UT	0	0	0	0	0	0	0	0	0	0	0 (0)	0	0
	WV	4	4	0	0	0	0	7	4	0	0	0	0 (0)	1
	WY	0	0	0	0	0	0	3	0	0	0	0	1	0 (0)
Total	**CT**	0 (0)	4	10	0	4	0	7	1	0	1	4	5	0
	HI	4	0 (0)	5	0	0	0	3	0	0	0	8	4	0
	ID	10	5	0 (0)	0	0	0	0	2	0	0	3	2	0
	MT	0	0	0	0 (0)	0	0	4	0	0	0	5	0	0
	NH	4	0	0	0	0 (0)	0	0	0	0	0	7	0	0
	ND	0	0	0	0	0	0 (0)	0	8	0	0	3	0	0

	State	CT	HI	ID	MT	NH	ND	OR	RI	SD	USVI	UT	WV	WY
	OR	8	3	0	4	0	0	0 (0)	7	0	0	8	8	3
	RI	1	0	2	0	0	8	7	0 (0)	4	0	2	6	0
	SD	0	0	0	0	0	0	0	4	0 (0)	0	3	0	0
	USVI	1	0	0	0	0	0	0	0	0	0 (0)	1	0	0
	UT	4	8	3	5	7	3	8	2	3	1	0 (0)	3	2
	WV	5	4	2	0	0	0	8	6	0	0	3	0 (0)	1
	WY	0	0	0	0	0	0	3	0	0	0	3	1	0 (0)

Following the administration, field-test items went through a substantial validation process. The process began with rubric validation. Rubric validation is a process in which a committee of state educators reviews student responses and the proposed scoring of those responses. The process is described in Volume 2, Section 2.7.1, Rubric Validation, of this technical report.

After rubric validation, classical item statistics were computed for the scoring assertions, including item difficulty and item discrimination statistics, testing time, and differential item functioning (DIF) statistics. The MOU established common standards for the statistics. Any items violating these standards were flagged for a second educator review. Even though the scoring assertions were the basic units of analysis used to compute classical item statistics, the business rules to flag items for another educator review were established at the item level because assertions cannot be reviewed in isolation. The statistics and business rules for flagging items are described in Section 4, Field-Test Classical Analysis. For each state, a data review committee consisting of educators (i.e., science teachers) supported by CAI content experts reviewed the items that were owned by the state and flagged for data review according to the established business rules. For ICCR, cross-state review committees were established.

Table 10 presents the number of field-test items administered in Montana, or another state or territory, the number of items rejected before or during rubric validation, the number of items sent for data review, and the number of items rejected during data review. The numbers in parentheses present the number of field-test items owned by Montana.

Table 10. Number of Field-Test Item Administration, Rubric Validation, and Item Data Review, Spring 2023

Grade Band and Item Type	Number of Field-Test Items Administered	Number of Items Rejected Before/During Rubric Validation	Number of Items Sent to Data Review	Number of Items Rejected at Data Review	Number of Items Remaining
Elementary School	**126 (11)**	**3 (0)**	**71 (7)**	**13 (1)**	**110 (10)**
Cluster	60 (4)	1 (0)	17 (0)	1 (0)	58 (4)
Stand-Alone	66 (7)	2 (0)	54 (7)	12 (1)	52 (6)
Middle School	**136 (9)**	**2 (0)**	**80 (5)**	**20 (2)**	**114 (7)**
Cluster	59 (5)	1 (0)	21 (2)	5 (1)	53 (4)
Stand-Alone	77 (4)	1 (0)	59 (3)	15 (1)	61 (3)
Total	**262 (20)**	**5 (0)**	**151 (12)**	**33 (3)**	**224 (17)**

Note. Montana-owned items are indicated in the parentheses.

Table 11 summarizes the Shared Science Assessment Item Bank after adding the field-test items that were administered in 2023 and passed rubric validation and item data review. The numbers in parentheses present the number of items owned by Montana.

Table 11. Summary of Shared Science Assessment Item Bank, Spring 2023

Grade Band and Item Type	Science Discipline			Item Bank Total[a]
	Earth and Space Sciences	*Life Sciences*	*Physical Sciences*	
Elementary School	**205 (12)**	**202 (14)**	**254 (14)**	**661 (40)**
Cluster	112 (5)	102 (4)	131 (4)	345 (13)
Stand-Alone	93 (7)	100 (10)	123 (10)	316 (27)
Middle School	**185 (12)**	**262 (10)**	**215 (9)**	**662 (31)**
Cluster	88 (4)	129 (5)	106 (3)	323 (12)
Stand-Alone	97 (8)	133 (5)	109 (6)	339 (19)
Total	**390 (24)**	**464 (24)**	**469 (23)**	**1323 (71)**

Note. Montana-owned items are indicated in the parentheses.
[a]Count excludes nine MOU items that do not align to the NGSS.

3.3 Test Design

The science tests were assembled under an adaptive test design, with the exception of the braille paper-pencil form. Tests were assembled using CAI's adaptive testing algorithm. The adaptive item selection algorithm selects items based on their content value and information value. At any given point during the test, the content value of an item is determined by its contribution to meeting the blueprint, given the content characteristics of the items that have already been administered. During the test, the content value increases for items that exhibit features that have not met their designated minimum as the end of the test approaches. Similarly, the content value decreases for items with content features for which the minimum has been met. The information value of an item is based on the item information function evaluated at the estimated proficiency. The proficiency estimate is updated throughout the test.

Under an adaptive test design, operational items are selected on the fly based on the performance of a student on past items while ensuring the test blueprint is followed for each individual student. The MSA blueprints are presented in this technical report in Volume 2, Section 4.2, Test Blueprints. Details of CAI's item selection algorithm are described in Volume 2, Appendix 2-L, Adaptive Algorithm Design.

The braille test, which is a paper-pencil test, was an accommodated fixed form. Form construction of the accommodated forms is discussed in Volume 2, Section 4.4, Paper-Pencil Accommodation Form Construction.

The main characteristics of the blueprint were that any performance expectation (PE) could be tested only once (indicated by the values of 0 and 1 for the minimum and maximum values of the individual PEs in the test blueprints; see Section 4.2, Test Blueprints of Volume 2). In general, no more than one item cluster or two stand-alone items could be sampled from the same Disciplinary Core Idea (DCI), and no more than three total items could be sampled from the same DCI (as indicated by the minimum and maximum values in the rows representing DCIs).

In the 2021 independent field test, a non-segmented test design was used. Students received items from different disciplines in random order. All the items administered were field-test items. Since 2021, a similar non-segmented test design with embedded field-test items was used. Embedded field-test items were randomly positioned in the test among operational items and randomly distributed across students. Every student received either one item cluster (average of 8.5 assertions per cluster) or four stand-alone items (average of 1.8 assertions per stand-alone) as field-test items throughout the test.

4. Field-Test Classical Analysis

As explained in Section 3, Item Bank and Test Design, science items administered as field-test items underwent rubric validation and data review. Items were flagged for data review based on business rules defined on classical item statistics. Except for response times, the classical item statistics are computed for individual assertions, whereas the business rules for flagging are defined at the item level. In general, item statistics used to flag items for data review were computed using the student responses of the state that owned the items. However, for Independent College and Career Readiness (ICCR) items, the flagging rules were defined on the item statistics computed from the combined data of states or territory that used ICCR items and administered either an independent or operational test. In 2023, those states were Connecticut, Idaho, New Hampshire, North Dakota, Oregon, Rhode Island, South Dakota, Utah, and West Virginia. Furthermore, to compute the differential item functioning (DIF) statistics for the field-test items, the data from all states with an operational or independent field test were combined to obtain a sufficient number of students for each demographic group. The criteria for flagging and reviewing items are provided in Table 12, and the statistics are described in Section 4.1, Item Discrimination, through Section 4.4, Differential Item Functioning Analysis. Items flagged for data review were reviewed by a committee, as explained in Section 3, Item Bank and Test Design.

Table 12. Thresholds for Flagging in Classical Item Analysis

Analysis Type	Flagging Criteria
Item Discrimination	Average biserial correlation < 0.25 (across the assertions within an item)
	One or more assertions with a biserial correlation < 0.05
Item Difficulty (Clusters)	Average *p*-value < 0.30 or > 0.85 (across the assertions within a cluster)
Item Difficulty (Stand-Alone Items)	Average *p*-value < 0.15 or > 0.95 (across the assertions within a stand-alone item)
Timing (Clusters)	Percentile 80* > 15 minutes
Timing (Stand-Alone Items)	Percentile 80* > 3 minutes
Timing	Assertions per minute < 0.5
DIF (Clusters)	Two or more assertions show “C” DIF in the same direction
DIF (Stand-Alone Items)	One or more assertions show “C” DIF

Note. *A percentile 80 of *x* minutes: 80% of the students spent *x* minutes or less on the item.

4.1 Item Discrimination

The item discrimination index indicates the extent to which each item differentiated between those test takers who possessed the skills being measured and those who did not. Generally, the higher the value, the better the item is able to differentiate between high- and low-achieving students.

For each assertion within an item, the discrimination index was calculated as the biserial correlation between the assertion score and the ability estimate for students. The average biserial correlation was then calculated across the assertions within an item.

4.2 Item Difficulty

Items that are either very difficult or very easy are flagged for review but are not necessarily removed from the item bank if they are grade-level appropriate and aligned with the test specifications. Both the *p*-value (item difficulty percent correct) for individual assertions and the average across all assertions of an item are calculated. Acceptable item *p*-values are summarized in Table 12.

4.3 Response Time

Given that the science item clusters consisted of multiple student interactions, they required more time for students to complete. Item response time was recorded and analyzed to ensure a good balance between the amount of information an item provided and the time students spent on the item. Specifically, the statistic "percentile 80" was computed for each item. A percentile 80 of *x* minutes means that 80% of the students spent *x* minutes or fewer on the item. An item was flagged for review when the

- percentile 80 > 15 minutes, if the item is an item cluster;
- percentile 80 > 3 minutes, if the item is a stand-alone item; or
- assertions per (percentile 80) minute < 0.5.

4.4 Differential Item Functioning

Differential item functioning (DIF) refers to items that appear to function differently across identifiable groups, typically across different demographic groups. Identifying DIF is important because it provides a statistical indicator that an item may contain cultural or other biases. DIF-flagged items are further examined by content experts who are asked to re-examine each flagged item to decide whether the item should be excluded from the pool due to bias. Not all items that exhibit DIF are biased, and various characteristics of the educational system may also lead to DIF.

CAI uses a generalized Mantel-Haenszel (MH) procedure to calculate DIF. The generalizations include adaptation to polytomous items and improved variance estimators to render the test statistics valid under complex sample designs. With this procedure, each student's estimated theta score on the operational items on a given test is used as the ability-matching variable. That score is divided into 10 intervals to compute the generalized MH chi-square ($GMH\chi^2$) DIF statistics for balancing the stability and sensitivity of the DIF scoring category selection. The standardized mean difference (SMD [Dorans & Schmitt, 1991]) was also computed.

The MH chi-square statistic (Holland & Thayer, 1988) is calculated as:

$$MH\chi^2 = \frac{(|\sum_k n_{R1k} - \sum_k E(n_{R1k})| - 0.5)^2}{\sum_k var(n_{R1k})},$$

where $k = \{1, 2, \dots K\}$ for the strata, $n_{R1k}$is the number of students with correct responses for the reference group in stratum k, and 0.5 is a continuity correction. The expected value is calculated as

$$E(n_{R1k}) = \frac{n_{+1k} n_{R+k}}{n_{++k}},$$

where n_{+1k} is the number of students with correct responses, n_{R+k} is the number of students in the reference group, and n_{++k} is the number of students in stratum k. The variance is calculated as

$$var(n_{R1k}) = \frac{n_{R+k} n_{F+k} n_{+1k} n_{+0k}}{n_{++k}^2 (n_{++k} - 1)},$$

where n_{F+k} is the number of students in the focal group, n_{+1k} is the number of students with correct responses, and n_{+0k} is the number of students with incorrect responses in stratum k.

The MH conditional odds ratio is calculated as

$$\alpha_{MH} = \frac{\sum_k n_{R1k} n_{F0k} / n_{++k}}{\sum_k n_{R0k} n_{F1k} / n_{++k}}.$$

The MH-delta (Δ_{MH} [Holland & Thayer, 1988]) is then defined as

$$\Delta_{MH} = -2.35 \ln(\alpha_{MH}).$$

The generalized MH statistic generalizes the MH statistic to polytomous items (Somes, 1986), and is defined as

$$GMH\chi^2 = \left(\sum_k \boldsymbol{a}_k - \sum_k E(\boldsymbol{a}_k)\right)' \left(\sum_k var(\boldsymbol{a}_k)\right)^{-1} \left(\sum_k \boldsymbol{a}_k - \sum_k E(\boldsymbol{a}_k)\right),$$

where $\boldsymbol{a}_k$ is a $(T-1) \times 1$ vector of item response scores and $E(\boldsymbol{a}_k)$ is a $(T-1) \times 1$ mean vector, both corresponding to the T response categories of a polytomous item (excluding one response); $var(\boldsymbol{a}_k)$ is a $(T-1) \times (T-1)$ covariance matrix calculated analogously to the corresponding elements in $MH\chi^2$ in stratum k.

The SMD (Dorans & Schmitt, 1991) is defined as

$$SMD = \sum_k p_{Fk} m_{Fk} - \sum_k p_{Fk} m_{Rk},$$

where

$$p_{Fk} = \frac{n_{F+k}}{n_{F++}}$$

is the proportion of the focal group students in stratum k,

$$m_{Fk} = \frac{1}{n_{F+k}} \left(\sum_t a_t n_{Ftk} \right)$$

is the mean item score for the focal group in stratum k, and

$$m_{Rk} = \frac{1}{n_{R+k}}\left(\sum_t a_t n_{Rtk}\right)$$

is the mean item score for the reference group in stratum k.

DIF analysis was conducted for all field-test items with at least 200 responses per item in each subgroup (Zwick, 2012) to detect potential item bias for major demographic groups. Student responses from multiple states were combined to minimize the number of items with insufficient sample sizes for one or more demographic groups.

DIF statistics were calculated at the assertion level and were performed for the following groups (some items had insufficient sample sizes for DIF analyses in some groups):

- Female vs. Male
- American Indian/Alaskan Native vs. White
- Asian vs. White
- African American vs. White
- Hawaiian/Pacific Islander vs. White
- Hispanic vs. White
- Multi-Racial vs. White
- English Learner (EL) vs. Non-EL
- Special Education (SPED) vs. Non-SPED
- Economically Disadvantaged vs. Non-Economically Disadvantaged

Similar to how the generalized MH procedure is used to classify items on traditional tests, assertions were classified into three categories (i.e., A, B, or C) for DIF, ranging from “no evidence of DIF” to “severe DIF.” The classification rules are shown in Table 13. Furthermore, assertions were categorized positively (i.e., +A, +B, or +C), signifying that an item favored the focal group (e.g., African American, Hispanic, female), or negatively (i.e., –A, –B, or –C), signifying that an item favored the reference group (e.g., White, male).

An item was flagged for data review according to the following criteria:

- **Item Clusters.** Two or more assertions showed “C” DIF in the same direction.
- **Stand-Alone Items.** One or more assertions showed “C” DIF.

Table 13. DIF Classification Rules

Assertions	
Category	*Rule*

C	MH_{χ^2} is significant and $\lvert SMD \rvert / \lvert SD \rvert \geq 0.25$.
B	MH_{χ^2} is significant and $\lvert SMD \rvert / \lvert SD \rvert < 0.25$.
A	MH_{χ^2} is not significant.

Note that, for the 2018 field test, a slightly less strict criterion was used for item clusters with 10 or more assertions (i.e., three or more assertions with "C" DIF in the same direction). The change was made taking into consideration the feedback received from several Technical Advisory Committees (TACs) and modified such that the rate of flagging items for DIF was similar for item clusters and stand-alone items (based on the flagging rates computed on items field-tested in 2018).

4.5 Classical Analysis Results

This section presents a summary of results from classical item analysis of the field-test items administered in 2023. A total of 20 field-test items were administered in Montana; 20 passed rubric validation. Among these 20 items, two were flagged for item discrimination, three were flagged for *p*-value, nine were flagged for response time, and none were flagged for DIF according to the criteria used in 2023 (as described in Section 4.1, Item Discrimination, through Section 4.4, Differential Item Functioning Analysis). Some items were flagged for multiple reasons. Flagged field-test items were reviewed by educators during data review. The total number of field-test items flagged and the total number of field-test items that passed item data review in 2023 are summarized in Table 10.

Table 14 through Table 16 provide the summary of the *p*-values, biserial correlations, and response times for the science field-test items administered in Montana in 2023 that passed rubric validation. The statistics were computed using Montana data only. For *p*-values and biserial correlations, the average values across the assertions within an item were used to compute percentiles and ranges.

Table 14. Distribution of p-Values for Field-Test Items, Spring 2023

Grade	Total FT Items	Min	5th Percentile	25th Percentile	50th Percentile	75th Percentile	95th Percentile	Max
5	11	0.26	0.28	0.33	0.43	0.57	0.64	0.68
8	9	0.11	0.13	0.20	0.33	0.36	0.48	0.49

Table 15. Distribution of Item Biserial Correlations for Field-Test Items, Spring 2023

Grade	Total FT Items	Min	5th Percentile	25th Percentile	50th Percentile	75th Percentile	95th Percentile	Max
5	11	0.24	0.33	0.45	0.50	0.54	0.65	0.69
8	9	0.33	0.34	0.40	0.44	0.52	0.54	0.54

Table 16. Summary of Response Times for Field-Test Items, Spring 2023

Grade	Item Type	Total FT Items	Min	5th Percentile	25th Percentile	50th Percentile	75th Percentile	95th Percentile	Max
5	Cluster	4	5.90	6.17	7.25	8.35	9.13	9.43	9.50
	Stand-Alone	7	2.20	2.29	2.55	2.80	3.35	4.09	4.30
8	Cluster	5	6.00	6.28	7.40	7.50	7.50	8.30	8.50
	Stand-Alone	4	1.50	1.74	2.70	3.20	3.55	4.15	4.30

Note, times are in minutes.

Table 17 presents the number of field-test items flagged for DIF for each item type and demographic group included in the DIF analyses in 2023.

Table 17. Differential Item Functioning Classifications for Field-Test Items, Spring 2023

DIF Flag	Item Type	Female/ Male	American Indian[a]/ White	Asian/ White	African American / White	Hawaiian[b] / White	Hispanic/ White	Multi-Racial/ White	EL/ Non-EL	SPED/ Non-SPED	Low Income/ Non-Low Income[c]
Grade 5											
Items Evaluated	Cluster	4	2	-	2	-	4	1	4	4	4
	Stand-Alone	7	3	-	3	-	7	-	7	7	7
Items Flagged C	Cluster	0	0	-	0	-	0	0	0	0	0
	Stand-Alone	0	0	-	0	-	0	-	0	0	0
Grade 8											
Items Evaluated	Cluster	5	2	-	-	-	5	-	5	5	5
	Stand-Alone	4	4	-	-	-	4	3	3	4	4
Items Flagged C	Cluster	0	0	-	-	-	0	-	0	0	0
	Stand-Alone	0	0	-	-	-	0	0	0	0	0

Note. 1. Full DIF group names: [a]American Indian/Alaskan Native; [b]Hawaiian/Pacific Islander; [c]Economically Disadvantaged vs. Non-Economically Disadvantaged
2. The symbol "-" means not enough sample size (less than 200 in either the reference or the focal group) to compute DIF statistics.

5. ITEM CALIBRATION

5.1 MODEL DESCRIPTION

In discussing item response theory (IRT) models for Montana, we distinguish between the underlying latent structure of a model and the parameterization of the item response function conditional on that assumed latent structure. Subsequently, we discuss how group effects are considered.

5.1.1 Latent Structure

Most operational assessment programs rely on a unidimensional IRT model for item calibration and computing scores for students. These models assume a single underlying trait and that items are independent given that underlying trait. In other words, the models assume that given the value of the underlying trait, knowing the response to one item provides no information about responses to other items. This assumption of conditional independence implies that the conditional probability of a pattern of I item responses takes the relatively simple form of a product over items for a single student, as shown below:

$$P(\boldsymbol{z_j}|\theta_j) = \prod_{i=1}^{I} P(z_{ij}|\theta_j), \qquad (1)$$

where z_{ij} represents the scored response of student j $(j = 1, \ldots, N)$ to item i $(i = 1, \ldots, I)$, $\boldsymbol{z}_j$ represents the pattern of scored item responses for student j, and θ_j represents student j's proficiency. Unidimensional IRT models differ with respect to the functional relation between the proficiency θ_j and the probability of obtaining a score z_{ij} on item i.

Montana Science Assessment items are more complex than traditional item types. A single item may contain multiple parts, and each part may contain multiple student interactions. For example, a student may be asked to select a term from a set of terms at several places in a single item. Instead of receiving a single score for each item, multiple inferences are made about the knowledge and skills that a student has demonstrated based on specific features of the student's responses to the item. These scoring units are called *assertions* and are the basic unit of analysis in our IRT analysis. That is, they fulfill the role of items in traditional assessments; however, for the Montana Science Assessment items, multiple assertions are typically developed around a single item so that assertions are clustered within items.

One approach is to apply one of the traditional IRT models to the scored assertions; however, a substantial complexity that arises from using this new item type is that local dependencies exist between assertions pertaining to the same stimulus (i.e., item or item cluster). The local dependencies between the assertions pertaining to the same stimulus constitute a violation of the assumption that a single latent trait can explain all dependencies between assertions. Fitting a unidimensional model in the presence of local dependencies may result in biased item parameters and standard errors of measurement (SEMs). In particular, it is well documented that ignoring local item dependencies leads to an overestimation of the amount of information conveyed by a set of responses and an underestimation of the SEM (e.g., Sireci, Thissen, & Wainer, 1991; Yen, 1993).

The effects of groups of assertions developed around a common stimulus can be accounted for by including additional dimensions corresponding to those groupings in the IRT model. These dimensions are considered to be nuisance dimensions[1]. Whereas traditional unidimensional IRT models assume that all assertions (the basic units of analysis) are independent given a single underlying trait θ, we now assume the conditional independence of assertions, given the underlying latent trait θ and all nuisance dimensions:

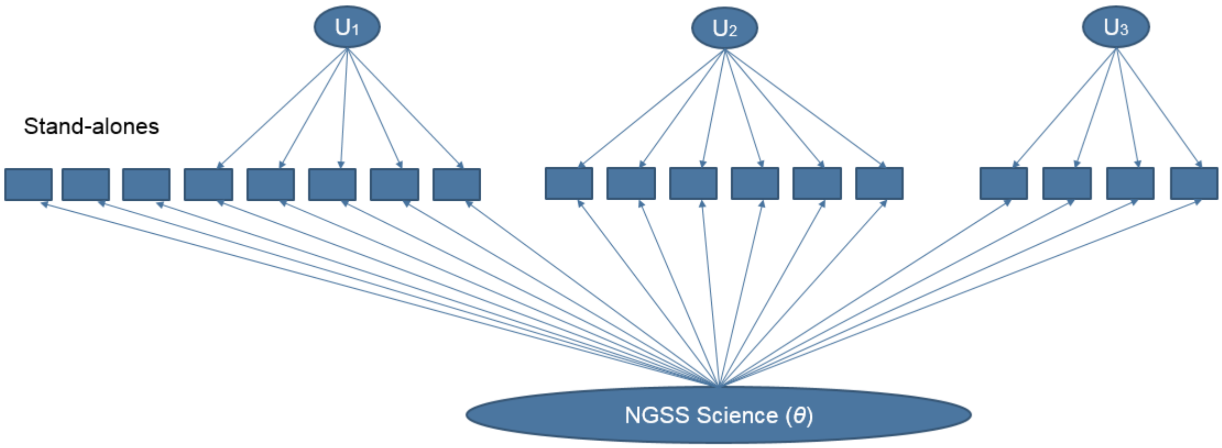
$$P\left(\mathbf{z}_{j} \middle| \theta_j, \mathbf{u}_j\right) = \prod_{i \in \text{SA}} P\left(z_{ij} \middle| \theta_j\right) \prod_{g=1}^{G} \prod_{i \in g} P\left(z_{ij} \middle| \theta_j, u_{jg}\right), \quad (2)$$

where SA indicates stand-alone item assertions, u_g indicates the nuisance dimension for assertion group g (with the position of student j on that dimension denoted as u_{jg}), and $\mathbf{u}$ is the vector of all G nuisance dimensions. It can be seen that the conditional probability $P\left(z_{ij} \middle| \theta_j, u_{jg}\right)$ becomes a function of two latent variables: the latent trait θ, representing a student's proficiency in science (the underlying trait of interest), and the nuisance dimension $u_{g,}$ accounting for the conditional dependencies between assertions of the same group. Ee assume that the nuisance dimensions are all uncorrelated with one another and with the general dimension. It is important to point out that even though every group of assertions introduces an additional dimension, models with this latent structure do not suffer from the complications of dimensionality like other multidimensional IRT models because one can take advantage of this special structure during model calibration (Gibbons & Hedeker, 1992). In this regard, Rijmen (2010) showed that it is unnecessary to assume all nuisance dimensions are uncorrelated; instead, it is sufficient that they are independent, given the general dimension θ.

The model structure of the IRT model for science is illustrated in Figure 1. Note that stand-alone items can be scored with more than one assertion. The assertions of stand-alone items with more than one assertion, but fewer than four assertions, are also modeled as stand-alone item assertions. Even though these assertions are likely to exhibit conditional dependencies, the variance of the nuisance dimension cannot be reliably estimated if it is based on a very small number of assertions. The few stand-alone items with four or more assertions are treated as item clusters to take into account the conditional dependencies.

[1] The term *nuisance dimension* pertains to within-item local dependencies among scoring assertions and should not be confused with the three dimensions of *A Framework for K–12 Science Education*.

Figure 1. Directed Graph of the Science IRT Model



5.1.2 Item Response Function

The item response functions of the stand-alone item assertions are modeled with a unidimensional model. For the grouped assertions, like in unidimensional models, different parametric forms can be assumed for the conditional probability of obtaining a score of z_{ij}. The Rasch testlet model (Wang & Wilson, 2005) is adopted as the IRT model for the Montana Science Assessment. For binary data, the Rasch testlet model is defined as:

$$P\left(z_{ij}|\theta_j, u_{jg}; b_i\right) = \frac{\exp\left(\theta_j + u_{jg} - b_i\right)}{1 + \exp\left(\theta_j + u_{jg} - b_i\right)}. \quad (3)$$

The item response function of the Rasch testlet model is the probability of a correct answer (i.e., a true assertion), as a function of the overall proficiency θ, the nuisance dimension u_g, and the item (i.e., assertion) difficulty b_i. The Rasch testlet model does not include item discrimination parameters; however, the same model structure as presented in Figure 1 could be employed with discrimination parameters included in Equations (2) and (3). Furthermore, only models for binary data are considered. Assertions are always binary because they are either true or false. Nevertheless, the model could easily accommodate polytomous responses by using the same response function incorporated in unidimensional models for polytomous data.

5.1.3 Multigroup Model

The Shared Science Assessment Item Bank is calibrated concurrently using all the items administered in any state that collaborates with CAI on their new science assessments. In the calibration, each state is treated as a population of students or a group. Overall group differences are taken into account by allowing a group-specific distribution of the overall proficiency variable

θ. Specifically, for every student j belonging to group k , $k = 1, \ldots, K$, a normal distribution is assumed,

$$\theta_j \sim N(\mu_k, \sigma_k^2),$$

where μ_k and σ_k^2 are the mean and variance of a normal distribution. The mean of the reference distribution ($k = 1$) is set to 0 to identify the model (for free item calibrations, where there are no anchor items with their location parameters set to specific values). For each of the nuisance variables u_g, a common variance parameter across groups is assumed, and the means are set to 0 in order to identify the model,

$$u_{jg} \sim N\left(0, \sigma_{u_g}^2\right).$$

5.2 Estimation

A separate IRT model is fit for each grade band. The parameters of the IRT model are estimated using the marginal maximum likelihood (MML) method. In the MML method, the latent proficiency variable θ_j and the vector of nuisance parameters $\boldsymbol{u}_j$ for each student j are treated as random effects and integrated out to obtain the marginal log likelihood corresponding to the observed response pattern $\boldsymbol{z}_j$ for student j,

$$\ell_j = \log \int \int P(\boldsymbol{z}_j | \theta_j, \boldsymbol{u}_j) N(\theta_j | \mu_k, \sigma_k^2) N(\boldsymbol{u}_j | \boldsymbol{0}, \boldsymbol{\Sigma}) d\boldsymbol{u}_j d\theta_j,$$

where $\boldsymbol{\Sigma}$ is a diagonal matrix with diagonal elements $\sigma_{u_k}^2$, denoting nuisance variance for group k. Across all students and groups, the overall log likelihood to be maximized with respect to the vector $\boldsymbol{\gamma}$ of all model parameters (i.e., item difficulty parameters and the mean and variance parameters of the latent variables) is

$$\ell(\boldsymbol{\gamma}) = \sum_k \sum_{j \in k} \ell_j.$$

Even though the number of latent variables in the overall log likelihood equation is very high, issues with dimensionality can be avoided because the integration over the high-dimensional latent $(\theta, \boldsymbol{u})$ space can be carried out as a sequence of computations in two-dimensional space $(\theta,\ \boldsymbol{u_g})$ (Gibbons & Hedeker, 1992; Rijmen, 2010).

The Shared Science Assessment Item Bank was calibrated freely in 2018 after the 2018 science test administrations concluded, and it was recalibrated in 2019 after the 2019 test administrations. Following 2019, field-test items are calibrated onto the scale of the Shared Science Assessment Item Bank by anchoring the operational items to their bank. In the anchored calibrations, the mean and variance of the overall science dimension are also estimated for each group.

Appendix 1-C, Calibration of the Shared Science Assessment Item Bank, contains a detailed description of the 2018 and 2019 calibration processes as well as a description of how the 2018 and 2019 scales were linked. Beginning 2024, CAI is beginning an item bank maintenance plan to evaluate bank performance since initial development.

Starting in 2021, CAIRT (Cambium Assessment IRT) is used to calibrate item parameters. CAIRT was specifically developed by CAI to calibrate the multigroup Rasch model on very large data sets because estimation times in commercially available software (i.e., flexMIRT) became prohibitive. CAIRT relies on the same estimation methods as the Bayesian networks with the logistic regression (BNL; Rijmen, 2006), a suite of Matlab functions for estimating a wide variety of latent variable models. BNL uses an efficient expectation-maximization (EM) algorithm based on graphical model theory (Rijmen, 2010). CAI has cross-validated parameter estimates from CAIRT with BNL and flexMIRT under various scenarios (Rijmen, Liao, & Lin, 2021). CAIRT is a web application that is available at no cost to members of the MOU.

Table 18 provides an overview of the groups per grade band for calibration of the 2023 field-test items. All but 81 items were calibrated on at least 1,500 student responses, and all but 5 items had a sample size larger than 1,200. The 5 items with fewer than 1,200 responses were either Oregon legacy items or Hawaii items in a research study.

Table 18. Groups Per Grade Band for the Spring 2023 Calibration of Field-Test Items

Group	Elementary School	Middle School
Connecticut	X	X
Hawaii	X	X
Idaho	X	X
Montana	X	X
New Hampshire	X	X
North Dakota	X	X
Oregon	X	X
Rhode Island	X	X
South Dakota	X	X
U.S. Virgin Islands		
Utah	X	X
West Virginia	X	X
Wyoming	X	X

5.2.1 Overview of the Operational Item Bank

Figure 2 and Figure 3 display the histogram of the difficulty parameters for grades 5 and 8 respectively, for all items that are part of the Montana Science Assessment operational pool. The figures also display the estimated student proficiency distributions for students who took the MSA. The distribution of the difficulty parameter overlaps well with the proficiency distribution in grade 5. The grade 8 items are slightly easier compared to the student proficiency distribution.

Figure 2. MSA Assertion Difficulty and Student Proficiency Distributions, Grade 5

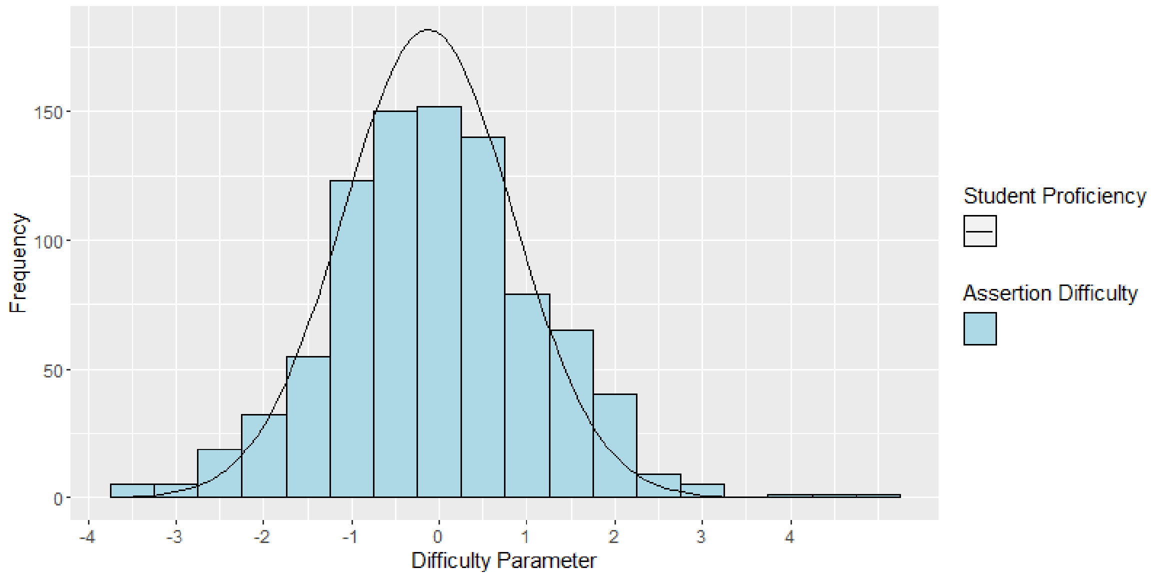
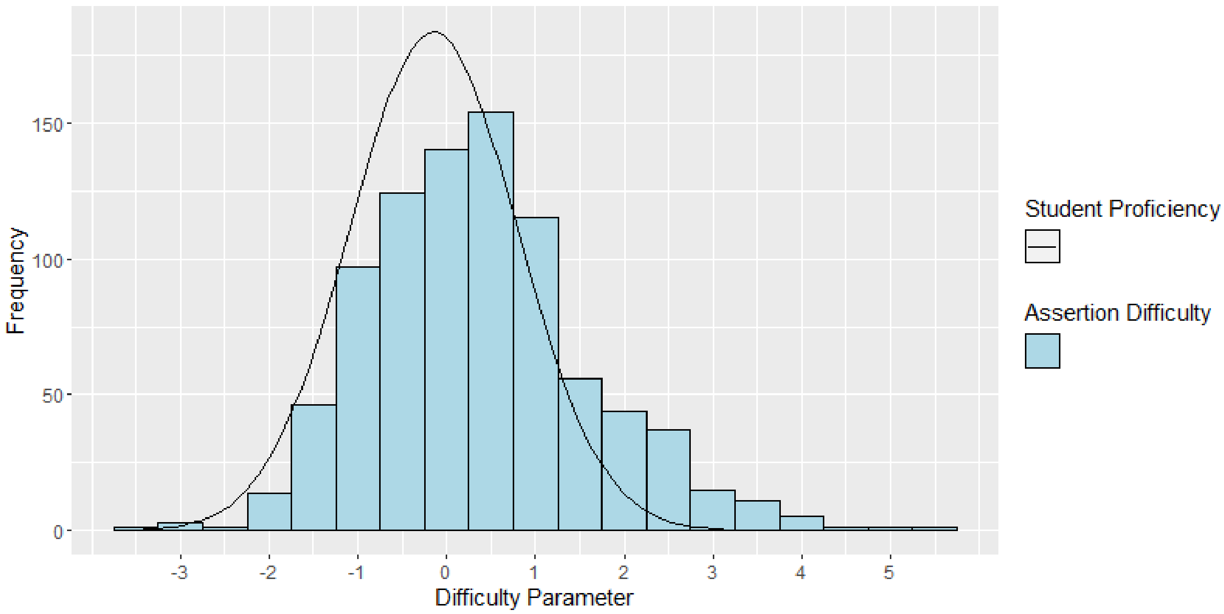


Figure 3. MSA Assertion Difficulty and Student Proficiency Distributions, Grade 8



6. SCORING

6.1 MARGINAL MAXIMUM LIKELIHOOD FUNCTION

Student scores are obtained by marginalizing out the nuisance dimensions $\boldsymbol{u}_j$ from the likelihood of the observed response pattern $\boldsymbol{z}_j$ for student j,

$$\ell_i(\theta_j) = log \int_{\boldsymbol{u}_j} P(\boldsymbol{z}_j|\theta_j, \boldsymbol{u}_j)\, N(\boldsymbol{u}_j|\boldsymbol{0}, \boldsymbol{\Sigma}) d\boldsymbol{u}_j,$$

and maximizing this marginalized likelihood function for θ_j. The marginal maximum likelihood estimation (MMLE) estimator is a hybrid between the expected a posteriori (EAP) estimator (by marginalizing out the nuisance dimensions) and the maximum likelihood estimation (MLE) estimator (by maximizing the resulting marginal likelihood for θ). The marginal likelihood is maximized with respect to θ using the Newton Raphson method. See Rijmen, Jiang, and Turhan (2018) for more details of the MMLE estimator and the validation study by Connecticut State Department of Education (2019) for the use of this estimator.

The proposed model reduces to the unidimensional Rasch model when the nuisance variances are zero for all g. Likewise, the proposed MMLE is equivalent to the MLE of the unidimensional Rasch model when all the nuisance variances are zero. This can be shown by using the variable transformation $\mathbf{v} = \Sigma^{-\frac{1}{2}}\mathbf{u}$. Then we have

$$\int_{\boldsymbol{u}_j} P(\boldsymbol{z}_j|\theta_j, \boldsymbol{u}_j)\, N(\boldsymbol{u}_j|\boldsymbol{0}, \boldsymbol{\Sigma}) d\boldsymbol{u}_j = \int_{\boldsymbol{v}_j} P\left(\boldsymbol{z}_j \middle|\theta_j, \boldsymbol{\Sigma}^{\frac{1}{2}}\boldsymbol{v}_j\right) N(\boldsymbol{v}_j|\boldsymbol{0}, \boldsymbol{I}) d\boldsymbol{v}_j.$$

If $\sigma^2_{u_g} = 0$ for all g, then

$$\int_{\boldsymbol{u}_j} P(\boldsymbol{z}_j|\theta_j, \boldsymbol{u}_j)\, N(\boldsymbol{u}_j|\boldsymbol{0}, \boldsymbol{\Sigma}) d\boldsymbol{u}_j = P(\boldsymbol{z}_j|\theta_j),$$

which is the likelihood under the unidimensional Rasch model.

6.2 DERIVATIVE

The marginal log likelihood function based on the item response theory (IRT) model with one overall dimension and one nuisance dimension for each grouping of assertions can be written as

$$l(\theta) = \sum_{i \in \mathrm{SA}} \log\big(P(z_i|\theta)\big) + \sum_{g=1}^{G} \log\left\{\int \mathrm{Exp}\left[\sum_{i \in g} \log\left(P(z_{ig}|\theta, u_g)\right)\right] N\left(u_g \middle| 0, \sigma^2_{u_g}\right) du_g\right\}.$$

The first derivative of the marginal log likelihood function with respect to θ is

$$\frac{dl(\theta)}{d\theta} = \sum_{i\in \mathrm{SA}} \frac{\frac{dP(z_i|\theta)}{d\theta}}{P(z_i|\theta)} + \sum_{g=1}^{G} \frac{\int \left\{ \mathrm{Exp}\left[\sum_{i\in g} \log\left(P\left(z_{ig}|\theta,u_g\right)\right)\right] \left(\sum_{i\in g} \frac{\frac{dP\left(z_{ig}|\theta,u_g\right)}{d\theta}}{P\left(z_{ig}|\theta,u_g\right)} \right) N\left(u_g\middle|0,\sigma_{u_g}^2\right) \right\} du_g}{\int \left\{ \mathrm{Exp}\left[\sum_{i\in g} \log\left(P\left(z_{ig}|\theta,u_g\right)\right)\right] N\left(u_g\middle|0,\sigma_{u_g}^2\right) \right\} du_g}$$

and the second derivative of the marginal log likelihood function with respect to θ is

$$\frac{d^2 l(\theta)}{d\theta^2} = \sum_{i\in \mathrm{SA}} \left[\frac{\frac{d^2 P(z_i|\theta)}{d\theta^2}}{P(z_i|\theta)} - \left(\frac{\frac{d\,P(z_i|\theta)}{d\theta}}{P(z_i|\theta)} \right)^2 \right]$$
$$+ \sum_{g=1}^{G} \frac{\int \mathrm{Exp}\left[\sum_{i\in g} \log\left(P\left(z_{ig}|\theta,u_g\right)\right)\right] \left(\sum_{i\in g} \frac{\frac{d\,P\left(z_{ig}|\theta,u_g\right)}{d\theta}}{P\left(z_{ig}|\theta,u_g\right)} \right)^2 N\left(u_g\middle|0,\sigma_{u_g}^2\right) du_g}{\int \left\{ \mathrm{Exp}\left[\sum_{i\in g} \log\left(P\left(z_{ig}|\theta,u_g\right)\right)\right] N\left(u_g\middle|0,\sigma_{u_g}^2\right) \right\} du_g}$$
$$+ \sum_{g=1}^{G} \frac{\int \mathrm{Exp}\left[\sum_{i\in g} \log\left(P\left(z_{ig}|\theta,u_g\right)\right)\right] \left(\sum_{i\in g} \left[\frac{\frac{d^2\,P\left(z_{ig}|\theta,u_g\right)}{d\theta^2}}{P\left(z_{ig}|\theta,u_g\right)} - \left(\frac{\frac{d\,P\left(z_{ig}|\theta,u_g\right)}{d\theta}}{P\left(z_{ig}|\theta,u_g\right)} \right)^2 \right] \right) N\left(u_g\middle|0,\sigma_{u_g}^2\right) du_g}{\int \left\{ \mathrm{Exp}\left[\sum_{i\in g} \log\left(P\left(z_{ig}|\theta,u_g\right)\right)\right] N\left(u_g\middle|0,\sigma_{u_g}^2\right) \right\} du_g}$$
$$- \sum_{g=1}^{G} \left\{ \frac{\int \mathrm{Exp}\left[\sum_{i\in g} \log\left(P\left(z_{ig}|\theta,u_g\right)\right)\right] \left(\sum_{i\in g} \frac{\frac{d\,P\left(z_{ig}|\theta,u_g\right)}{d\theta}}{P\left(z_{ig}|\theta,u_g\right)} \right) N\left(u_g\middle|0,\sigma_{u_g}^2\right) du_g}{\int \left\{ \mathrm{Exp}\left[\sum_{i\in g} \log\left(P\left(z_{ig}|\theta,u_g\right)\right)\right] N\left(u_g\middle|0,\sigma_{u_g}^2\right) \right\} du_g} \right\}^2 .$$

Based on the above equations, we only need to define the ratios of the first and second derivatives of the item response probabilities with respect to θ to the response probabilities. For the Rasch testlet model, these are obtained as

$$p_i = P(z_i = 1|\theta) = \frac{\mathrm{Exp}(\theta - b_i)}{1 + E\mathrm{xp}(\theta - b_i)},\ q_i = P(z_i = 0|\theta) = 1 - p_i,$$

and

$$p_{ig} = P(z_{ig} = 1|\theta, u_g) = \frac{Exp(\theta + u_g - b_i)}{1 + Exp(\theta + u_g - b_i)},\ q_{ig} = P(z_{ig} = 0|\theta, u_g) = 1 - p_{ig}.$$

Therefore, we have,

$$\frac{\frac{dp_i}{d\theta}}{p_i} = q_i\ ,\ \frac{\frac{dq_i}{d\theta}}{q_i} = -p_i,$$

$$\frac{\frac{dp_{ig}}{d\theta}}{p_{ig}} = q_{ig}\ ,\ \frac{\frac{dq_{ig}}{d\theta}}{q_{ig}} = -p_{ig},$$

$$\frac{\frac{d^2 p_i}{d\theta^2}}{p_i} - \left(\frac{\frac{d\,p_i}{d\theta}}{p_i}\right)^2 = -p_i q_i,$$

$$\frac{\frac{d^2 q_i}{d\theta^2}}{q_i} - \left(\frac{\frac{d\,q_i}{d\theta}}{q_i}\right)^2 = -p_i q_i,$$

$$\frac{\frac{d^2 p_{ig}}{d\theta^2}}{p_{ig}} - \left(\frac{\frac{d\,p_{ig}}{d\theta}}{p_{ig}}\right)^2 = -p_{ig} q_{ig}, \text{ and}$$

$$\frac{\frac{d^2 q_{ig}}{d\theta^2}}{q_{ig}} - \left(\frac{\frac{d\,q_{ig}}{d\theta}}{q_{ig}}\right)^2 = -p_{ig} q_{ig}.$$

6.3 Extreme Case Handling

As with the MLE, the MMLE is not defined for zero and perfect scores. These cases are handled by assigning the lowest obtainable theta (LOT) scores and highest obtainable theta (HOT) scores, respectively. Table 19 contains the LOT and HOT values for each grade.

6.4 Standard Error of Measurement

The standard error of measurement (SEM) of the MMLE score estimate is:

$$SEM(\hat{\theta}_{MMLE}) = \frac{1}{\sqrt{I(\hat{\theta}_{MMLE})}}$$

where $I(\hat{\theta}_{MMLE})$ is the observed information evaluated at $\hat{\theta}_{MMLE}$. The observed information is calculated as $I(\theta^2) = -\frac{d^2 l(\theta)}{d\theta^2}$, where $\frac{d^2 l(\theta)}{d\theta^2}$ is defined in Section 6.2, Derivative. Note that the calculation of the SEM depends on the unique set of items that each student answers and their estimate of θ. Different students have different SEM values, even if they have the same raw score and/or theta estimate. Standard errors are truncated at 1 for the overall science scores and truncated at 1.4 for the discipline scores.

Standard errors for MMLE estimates truncated at the LOT and HOT are computed by evaluating the observed information at the MMLE before truncation. For all incorrect or all correct answers, the reported SEM is set at the truncation value for the standard error.

6.5 Scoring Incomplete Tests

For science, a test is considered “attempted” if a student responded to at least one item (cluster or stand-alone). An attempted test is considered complete if the student responds to all the operational items. Otherwise, the test is “incomplete.”

Tests that are attempted but incomplete receive overall science scores. In order to receive a discipline score (e.g., Life Sciences, Physical Sciences, Earth and Space Sciences), a student must have attempted the corresponding discipline of the test. The MMLE is used to score the attempted incomplete tests, counting unanswered items as incorrect. If the identities of the unanswered items are unknown due to the test being assembled on the fly, the item parameters for a “typical” item are used. If a missing item is an item cluster, the simulated item parameters of the missing item are the item parameters of item cluster 637 for grade 5 and item cluster 11 for grade 8, which are operational item clusters that are typical for the Montana Science Assessment item pool used in Montana in terms of the number of assertions and estimated parameters. Likewise, if a missing item is a stand-alone item, the simulated item parameters of the missing item are the item parameters of stand-alone item 671 for grade 5 and item 642 for grade 8, which are operational stand-alone items that are typical for the Montana Science Assessment item pool used in Montana.

If the identities of items that have not been answered are known because they have already been lined up through the pre-fetch process, the item parameters of the lined-up items will be used. Similarly, for the accommodated forms that are fixed-forms, the item parameters of the unanswered items on the form will be used.

6.6 Student-Level Scale Score

At the student level, scale scores are computed for

1. Overall Science
2. Life Sciences
3. Physical Sciences
4. Earth and Space Sciences

Scores are computed using the MMLE method outlined in this report, with all items from overall science or only items within the given discipline. Scores are truncated on the “theta” scale at the LOT and HOT values specified in Table 19, which correspond to values of the estimated mean minus/plus four times the estimated standard deviation of θ.

The reporting scales will be a linear transformation of the theta scales

$$SS = a * \hat{\theta}_{MMLE} + b,$$

where a and b are the slope and intercept of the linear transformation that transforms $\hat{\theta}_{MMLE}$ to the reporting scale (refer to Table 19). The SEM for the estimated scale score is obtained as

$$SEM_{SS} = a * SEM_{\hat{\theta}_{MMLE}}.$$

In 2022, slope a and intercept b were chosen so that the center of the reporting scale of each grade (500 and 800, respectively) is at the grade mean of the 2022 base-year and has a standard deviation of 25. Specifically, for grade 5, the slope and intercept were obtained as

$$\begin{aligned} SS &= 25\theta^* + 500 \\ &= 25\frac{\theta - \hat{\mu}_\theta}{\hat{\sigma}_\theta} + 500 \\ &= \frac{25}{\hat{\sigma}_\theta}\theta + \left(500 - \frac{25\hat{\mu}_\theta}{\hat{\sigma}_\theta}\right), \end{aligned}$$

where the second line stems from standardizing theta, $\theta^* = \frac{\theta - \hat{\mu}_\theta}{\hat{\sigma}_\theta}$. For grades 8, the slope and intercept can also be derived similarly.

Per grade, Table 19 presents the intercept, slope, LOT, HOT, lowest obtainable scale score (LOSS), and highest obtainable scale score (HOSS) values used for the 2023 reporting scale. The scale score distribution is reported for overall science in Appendix 1-D, Distribution of Scale Scores and Performance Levels. The scale score distribution is reported for the science disciplines in Appendix 1-E, Distribution of Scale Scores by Science Discipline.

Table 19. Science Reporting Scale Linear Transformation Constants, Theta, and Corresponding Scaled-Score Limits for Extreme Ability Estimates

Grade	Slope (*a*)	Intercept (*b*)	Lowest of Theta (LOT)	Highest of Theta (HOT)	Lowest of Scale Score (LOSS)	Highest of Scale Score (HOSS)
5	26.381	502.218	-3.87	3.70	400	600
8	27.284	804.602	-3.83	3.49	700	900

6.7 Rules for Calculating Performance Levels

Performance levels and corresponding cut scores were set during standard setting in summer 2022. Students are classified into one of four performance levels, based on their total score. The distribution of performance levels is summarized in Appendix 1-D, Distribution of Scale Scores and Performance Levels. Further, the distribution of scale scores and performance levels for subgroups described in Section 4.4, Differential Item Functioning, are presented in Appendix 1-F, Distribution of Scale Scores and Performance Levels by Subgroup.

Table 20 lists the cut scores on the reporting scale metrics for each grade.

Table 20. Performance-Level Cut Scores

Grade	Cut 1	Cut 2	Cut 3
5	477	506	531
8	781	808	826

6.7.1 Strengths and Weaknesses for Disciplines Relative to Proficiency Cut Score

Discipline-level classifications are computed to classify student performance levels for each of the science disciplines/areas of science. The following are the classification rules:

- If $\hat{\theta}_{discipline} < \theta_{proficient} - 1.5 * SEM\left(\hat{\theta}_{discipline}\right)$, performance is classified as *Below Standard*;
- if $\theta_{proficient} - 1.5 * SEM\left(\hat{\theta}_{discipline}\right) \leq \hat{\theta}_{discipline} < \theta_{proficient} + 1.5 * SEM\left(\hat{\theta}_{discipline}\right)$, performance is classified as *Approaching Standard*; and
- if $\hat{\theta}_{discipline} \geq \theta_{proficient} + 1.5 * SEM\left(\hat{\theta}_{discipline}\right)$, performance is classified as *Above Standard*,

where $\theta_{Proficient}$ is the proficiency cut score of the overall test. Standard errors are truncated at 1.4. The LOT is always classified as *Below Standard*, and the HOT is always classified as *Above Standard*.

6.8 Residual-Based Reporting at the Level of Disciplinary Core Ideas

6.8.1 Relative to Overall Performance

For aggregated units (i.e., classrooms, schools, and districts), there is residual-based reporting at more fine-grained levels. In 2022, reports were provided at the level of Disciplinary Core Ideas (DCIs).

For each assertion i, the residual between the observed and expected score for each student j is defined as

$$\delta_{ij} = z_{ij} - E\left(z_{ij}\right).$$

The expected score is computed for a student's estimated overall ability. For the assertions clustered within an item, the expected score is marginalized over the nuisance dimensions for the assertions clustered within an item,

$$E\left(z_{ijg} = 1; \theta_{j,overall}, \boldsymbol{\tau}_i\right) = \int P\left(z_{ijg} = 1 | u_{jg};\ \theta_{j,overall}, \boldsymbol{\tau}_i\right) N\left(u_{jg}\right) du_{jg},$$

where $\boldsymbol{\tau}_i$ is the vector of parameters for assertion i (e.g., for the Rasch testlet model, $\boldsymbol{\tau}_i = b_i$), and $P\left(z_{ijg} = 1 | u_{jg};\ \theta_{j,overall}, \boldsymbol{\tau}_i\right)$ is defined in Section 6.2, Derivative. Next, residuals are aggregated over assertions within each student,

$$\delta_{jDCI} = \frac{\sum_{i \in DCI} \delta_{ij}}{n_{jDCI}},$$

and over students of the group on which is reported,

$$\bar{\delta}_{DCIm} = \frac{1}{n_m} \sum_{j \in m} \delta_{jDCI},$$

where n_{jDCI} is the number of assertions related to the DCI for student j, and n_m is the number of students in a group assessed on the DCI. If a student did not see any items on a DCI, the student is not included in the n_m count for the aggregate. The standard error of the average residual is computed as

$$SEM(\bar{\delta}_{DCIm}) = \sqrt{\frac{1}{n_m(n_m - 1)} \sum_{j \in m} \left(\delta_{jDCI} - \bar{\delta}_{DCIm}\right)^2}.$$

A statistically significant difference from zero in these aggregates is evidence that a class, teacher, school, or district is more effective (if $\bar{\delta}_{DCIm}$ is positive) or less effective (negative $\bar{\delta}_{DCIm}$) in teaching a given DCI.

We do not suggest direct reporting of the statistic $\bar{\delta}_{DCIm}$; instead, we recommend reporting in the aggregate whether a group of students performs better, worse, or as expected on this DCI. It will also be indicated that, in some cases, sufficient information is not available.

For target-level strengths/weakness, the following is reported:

- If $\bar{\delta}_{DCIm} \leq -1.5 * SEM\left(\bar{\delta}_{DCIm}\right)$, then performance is *worse than* on the overall test.
- If $\bar{\delta}_{DCIm} \geq 1.5 * SEM(\bar{\delta}_{DCIm})$, then performance is *better than* on the overall test.
- Otherwise, performance is *similar to* on the overall test.
- If $SEM\left(\bar{\delta}_{DCIm}\right) > 0.2$, data are insufficient.

6.8.2 Relative to Proficiency Cut Score

DCI-level scores for aggregated units can be computed using the same method as outlined in Section 6.8.1, Relative to Overall Performance, but with the expected score computed at the theta value corresponding to the proficiency cut score:

$$E\left(z_{ijg} = 1; \theta_{proficiency}, \boldsymbol{\tau}_i\right) = \int P\left(z_{ijg} = 1 | u_{jg};\ \theta_{proficiency}, \boldsymbol{\tau}_i\right) N\left(u_{jg}\right) du_{jg}.$$

The following is reported for DCIs for aggregate units:

- If $\bar{\delta}_{DCIm} \leq -\ 1.5 * SEM\left(\bar{\delta}_{DCIm}\right)$, then performance is *below* the proficiency cut score.
- If $\bar{\delta}_{DCIm} \geq 1.5 *\ SEM\left(\bar{\delta}_{DCIm}\right)$, then performance is *above* the proficiency cut score.

- Otherwise, performance is *approaching* the proficiency cut score.
- If $SEM(\bar{\delta}_{DCIm}) > 0.2$, data are insufficient.

7. QUALITY CONTROL PROCEDURES

Cambium Assessment, Inc.'s (CAI) quality assurance (QA) procedures are built on two key principles: automation and replication. Certain procedures can be automated (e.g., Blueprint Match, Section 7.1.2 of this volume), which removes the potential for human error. Procedures that cannot be reasonably automated are replicated by two independent analysts at CAI.

Although the quality of any test is monitored as an ongoing activity, several sources of CAI's quality control system are described here. First, QA reports are routinely generated and evaluated throughout the testing window to ensure that each test performs as anticipated. Second, the quality of scores is ensured by employing a second independent scoring verification system.

7.1 QUALITY ASSURANCE REPORTS

Test monitoring occurs while tests are administered in a live environment to ensure that item behavior is consistent with expectations. This is accomplished using CAI's Quality Assurance (QA) System that yields item statistics, blueprint match rates, and item exposure rate reports.

7.1.1 Item Analysis

The item analysis report is a key check for the early detection of potential problems with item scoring, including the incorrect designation of a keyed response or other scoring errors and potential breaches of test security that may be indicated by changes in the difficulty of test items. To examine the performance of test items, this report generates classical item analysis indicators of difficulty and discrimination, including proportion correct, biserial/polyserial correlation, and item fit statistics based on the item response theory model. The report is configurable and can be produced to flag only items with statistics that fall outside a specified range or to generate reports based on all items in the pool. For science, statistics reports at the assertion level (which are the units of analysis for science) are not yet available; however, as a routine and continuing practice, CAI psychometricians compute and monitor classical item statistics and item fit (i.e., item drift) at the end of each testing window.

7.1.2 Blueprint Match

The QA system generates Blueprint Match reports at the content-standards level and for other content requirements, such as strand and affinity group for science. For each blueprint element, the report indicates the minimum and maximum number of items specified in the blueprint, the number of test administrations in which those specifications were met, the number of administrations in which the blueprint requirements were not met, and, for administrations in which specifications were not met, the number of items by which the requirement was not met.

In Spring 2023, no blueprint violations were found in both grades across all test events. Blueprint match is discussed in detail in this technical report in Volume 2, Test Development, for both simulated and operational test administrations.

7.1.3 Item Exposure Rates

The QA system also generates item exposure reports that allow test items to be monitored for unexpectedly large exposure rates or unusually low item-pool usage throughout the testing window. As with other reports, it is possible to examine the exposure rate for all items or flagged items with exposure rates that exceed an acceptable range. Often, item overexposure indicates a blueprint element or combination of blueprint elements that are underrepresented in the item pool and should be targeted for future item development. Such item overexposure is also usually anticipated in the simulation studies used to configure the adaptive algorithm. Details about the item exposure rates in both the simulation studies and the operational data are discussed in Volume 2, Test Development, of this technical report.

7.1.4 Cheating Detection Analysis

The CAI QA system can also provide a forensics report to identify possible irregularities in test administration for further investigation. Unusual patterns of responding at the student level can be aggregated to the test session, test administrator, and school levels to identify possible group-level testing anomalies. CAI psychometricians can monitor testing anomalies throughout the testing window. Evidence can be evaluated with respect to item response times, and irregular item response patterns using the person-fit index. The flagging criteria used for these analyses are configurable and can be changed by the user. The analyses used to detect the testing anomalies can be run anytime within the testing window.

7.2 Scoring Quality Check

All student test scores are produced using CAI's scoring engine. Before releasing any scores, a second score verification system is used to verify that all test scores match with 100% agreement in all tested grades. The second system is independently constructed and maintained from the main scoring engine and separately estimates scores using the procedures described within this report.

8. References

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

Cai, L. (2017). flexMIRT®: Flexible multilevel multidimensional item analysis and test scoring (version 3.51) [computer software]. Chapel Hill, NC: Vector Psychometric Group.

Dorans, N. J., & Schmitt, A. P. (1991). *Constructed response and differential item functioning: A pragmatic approach* (ETS Research Report No. 91–47). Princeton, NJ: Educational Testing Service.

Gibbons, R. D., & Hedeker, D. R. (1992). Full-information item bi-factor analysis. *Psychometrika*, *57*(3), 423–436.

Holland, P. W., & Thayer, D. T. (1988). Differential item functioning and the Mantel-Haenszel procedure. In H. Wainer & H. I. Braun (Eds.), *Test validity* (pp. 129–145). Hillsdale, NJ: Lawrence Erlbaum Associates.

Koretz, D., & Hamilton, L. S. (2006). Testing for accountability in K–12. In R. L. Brennan (Ed.), *Educational measurement* (4th ed., pp. 531–578). Westport, CT: American Council on Education/Praeger.

National Research Council. (2012). *A framework for K–12 science education: Practices, crosscutting concepts, and core ideas*. Washington, DC: The National Academies Press.

Rijmen, F. (2006). *BNL: A Matlab toolbox for Bayesian networks with logistic regression nodes* (Technical Report). Amsterdam: VU University Medical Center.

Rijmen, F. (2010). Formal relations and empirical comparison among the bi-factor, the testlet, and a second-order multidimensional IRT model. *Journal of Educational Measurement*, *47*(3), 361–372.

Rijmen, F., Jiang, T., & Turhan, A. (2018, April). *An item response theory model for new science assessments*. Paper presented at the National Council on Measurement in Education, New York, NY.Rijmen, F., Liao, D., & Lin, Z. (2021). *The Rasch testlet model for the calibration of three-dimensional science assessments: A software comparison* [White paper]. Washington, DC: Cambium Assessment, Inc.

Sireci, S. G., Thissen, D., & Wainer, H. (1991). On the reliability of testlet-based tests. *Journal of Educational Measurement, 28*, 237–247.

Somes, G. W. (1986). The generalized Mantel Haenszel statistic. *The American Statistician*, *40*, 106–108.

Wang, W. C., & Wilson, M. (2005). The Rasch testlet model. *Applied Psychological Measurement*, *29*(2), 126–149.

Yen, W. (1993). Scaling performance assessments: Strategies for managing local item dependence. *Journal of Educational Measurement*, *30*, 187–213.

Zwick, R. (2012). *A review of ETS differential item functioning assessment procedures: Flagging rules, minimum sample size requirements, and criterion refinement* (ETS Research Report No. 12–08). Princeton, NJ: Educational Testing Service.

Appendix 1-A

Caveon Test Security Overview

Caveon Test Security Overview

Test Administration Security – Caveon

The Cambium Assessment, Inc (CAI) utilizes the Caveon Web Patrol™ service to support test security compliance. Caveon is recognized as the only full-service test security organization that has national experience and expertise in this area, and its Web Patrol™ services have been used for Virginia, Maryland, Mississippi, and PARCC assessments. By scouring the Internet and public-facing social media sites for breaches in test security, Caveon can systematically find and track threats to the testing program.

Web Patrol leverages the best of both automated technologies and our human capacity to judge and analyze. The result of this unique combination is a service that continually and systematically finds and tracks threats to the testing program.

Description

Caveon Web Patrol leverages technology tools and human expertise to identify, prioritize, and monitor websites, discussion forums, peer-to-peer servers, etc., where sensitive test information may be disclosed or at risk of disclosure.

Patrolling efforts routinely find and evaluate “brain-dumps” (websites where test questions have been posted, supposedly by individuals who memorized them and/or where disclosed test content may be inexpensively resold), test preparation training/education sites that may use actual (operational) test questions in the training, online auctions and classifieds such as eBay and Craigslist, and other social media channels and forums in which actual test items may be revealed or proxy test-takers offer their services. Regular update reports are generated that categorize identified threats by level of actual or potential risk to the testing program based on the representations made on the websites, or actual analysis of the proffered content. Websites and Internet extracts are ranked from CLEARED (Lowest risk but should be monitored) to SEVERE (Highest risk). The reports contain specific URLs and other content extractions that represent and depict the categorized threat. Additionally, the reports include overall and specific threat analysis, with actionable recommendations as well as any anticipated mitigation strategies for the Montana Office of Public Instruction to follow in minimizing and removing the dangers.

Comprehensive, Consistent Monitoring

In conducting web patrol operations, Caveon utilizes a team of specialists who spend days and evenings continually trolling the Internet for intellectual property, the team leverages numerous search technologies, some licensed and some publicly accessible (e.g., “Open Source”), to ensure comprehensive, consistent, and continual monitoring of the web.

Verifying and Managing Threats

Casting such a broad net across the web means the team must cull through thousands of search results (each is a possible threat) and dig deeper to explore whether a result is benign or a legitimate worry. Team members have, after years of service, become experts at quickly reviewing a search hit and discerning a level of risk. Despite technology innovations in other aspects of the service, this work requires human judgment and is vitally necessary to take action against real threats to test security.

Once a threat is verified, CAI and Caveon coordinate with Montana Office of Public Instruction to systematically work through the steps necessary to have infringing content removed.

An escalation path of legal remedies is available. That path begins with formal “bystander” notifications and cease and desist letters. The path ends when the website operators remove copyrighted material and/or cease operations, either voluntarily or by compulsion. CAI endeavors to complement existing activities of Montana Office of Public Instruction, including issuing formal notices under existing U.S. copyright laws to offending website owners, ISPs, search engines, etc. Keys to successful threat removal include the following:

Timeliness of Notification

By continually, systematically patrolling for new threats and monitoring existing ones, Caveon Web Patrol quickly ascertains when a breach has or may occur. When a breach has been discovered, CAI will immediately notify the Montana Office of Public Instruction.

Assistance Taking Down Material

Immediate notification of dangerous threats to the testing program is only half the solution. With direction and support from the Montana Office of Public Instruction, CAI provides quick front line support through various means to take the next step, neutralizing the hazard. There are multiple options at Montana Office of Public Instruction’s disposal to help protect its IP. CAI has experience with:

- Bystander Letters

Bystander letters can be sent immediately to website operators upon threat detection by Caveon. In most cases, simply alerting operators that copyrighted materials may be published on their websites is enough to get it removed.

Caveon Web Patrol service begins one week prior to the opening of the administration window and continues for one week after the test administration window closes.

Appendix 1-B

Shared Science Assessment Item Bank: Field-Testing

TABLE OF CONTENTS

List of Tables

The Shared Science Assessment Item Bank is the product of a collaboration between Cambium Assessment, Inc. (CAI), multiple states and one U.S. territory that share a Memorandum of Understanding (MOU). Every participant of the MOU contributes items to the Shared Science Assessment Item Bank that underwent the same development process. The portion of the bank contributed by CAI is part of the Independent College and Career Readiness (ICCR) bank. Since the start of the 2017–2018 school year, items have been field-tested annually. This appendix describes how the field tests were conducted from 2017–2018 through 2021–2022 across the states that rely on the Shared Science Assessment Item Bank, or the ICCR portion thereof, for their three-dimensional science assessments.

1. 2018 Field Tests

In 2018, a large pool of items was field-tested in nine states. For three states (Hawaii, Oregon, and Wyoming), unscored field-test items were added as an additional segment to the operational (scored) legacy science test. Two other states (Connecticut and Rhode Island) conducted an independent field test in which all students participated and were administered a full set of items, but no scores were reported. In the remaining four states that field-tested items from the Shared Science Assessment Item Bank (New Hampshire, Utah, Vermont, and West Virginia), an operational field test was administered, meaning tests consisted of scored field-test items. Items became operational and were scored after the test administration if they were not rejected during rubric validation or item data review, as described later in this section. In total, 340 item clusters and 205 stand-alone items were administered in the elementary, middle, and high school grade bands. Table 1 presents the number of item clusters and stand-alone items administered in each grade band for each state.

Table 1. Number of Field-Test Items Administered, Spring 2018

Grade Band and Item Type	CT	HI	MSSA[a]	NH	OR	UT	WV	WY	Total
Elementary School	**135**	**24**	**69**	**58**	**26**	**–**	**91**	**14**	**153**
Cluster	78	13	40	34	20	–	56	6	86
Stand-Alone	57	11	29	24	6	–	35	8	67
Middle School	**174**	**27**	**56**	**55**	**28**	**98**	**123**	**17**	**241**
Cluster	115	13	26	30	22	98	90	5	171
Stand-Alone	59	14	30	25	6	–	33	12	70
Total	**309**	**51**	**125**	**113**	**54**	**98**	**214**	**31**	**394**

Note. [a]MSSA = Rhode Island and Vermont's Multi-State Science Assessment

For the states with a separate field-test segment (states with a legacy science test) and one of the states with an operational field test (Utah), fixed field-test forms were constructed (using a balanced incomplete design for all states excepting Utah) and randomly assigned so that the group of students administered one form was comparable to the groups of students that were assigned other forms.

For the independent and operational field tests (excepting in Utah), items were administered using a linear-on-the-fly (LOFT) test design in which items are selected on the fly, resulting in a unique test form for each student. The difference between the test design for the independent field tests and operational field tests depended on the test blueprint. The only blueprint constraint imposed on the independent field tests was that students received four stand-alone items and two item clusters for each of the three science disciplines. In contrast, a full blueprint was implemented for the states with an operational field test. In contrast, a full blueprint was implemented for the states with an operational field test.

There was a target of a minimum sample size of 1,500 students per item for any given state. Most items were administered in two or more states so that the item pools for all individual states were linked through common items. Approximately 98.3% of the items met or exceeded the target sample size of 1,500 in at least one state, while 98.8% of the items had a sample size of at least 1,350 (10% of the target) in at least one state. The common item design was used to calibrate all the items on a common science scale.

Table 2 and Table 3 present the number of item clusters and stand-alone items that were common between the item pools of any two states. The numbers below the shaded cells represent the number of common field-test items between any two states, and the numbers above the shaded cells represent the number of common items that survived rubric validation and were included in the 2018 calibration. In each of the shaded cells, the number outside the parentheses represents the number of unique items administered only in the given state, and the number provided in parentheses represents the number of unique and/or common items that were calibrated with the data from that state only. Table 2 presents the results for elementary school, and Table 3 presents the results for middle school. The numbers at field-testing differ slightly from the numbers at calibration for various reasons, such as items not passing rubric validation and versioning issues for some items in some states.

Table 2. Number of Common Elementary School Field-Test Items Administered and Calibrated, Spring 2018

	State	CT	HI	MSSA[a]	NH	OR	UT	WV	WY
Cluster	CT	**3 (3)**	9	36	28	16	–	49	6
	HI	10	**0 (0)**	7	8	5	–	12	1
	MSSA	36	8	**0 (2)**	15	12	–	26	2
	NH	30	8	17	**1 (3)**	5	–	22	2
	OR	17	5	13	5	**1 (1)**	–	5	1
	UT	–	–	–	–	–	**–**	–	–
	WV	49	12	27	25	5	–	**0 (4)**	2
	WY	6	1	2	2	1	–	2	**0 (0)**
Stand-Alone	CT	**1 (3)**	5	25	22	2	–	33	7
	HI	5	**6 (6)**	0	0	0	–	4	0

	State	CT	HI	MSSA[a]	NH	OR	UT	WV	WY
	MSSA	26	0	**0 (1)**	10	4	–	13	3
	NH	24	0	11	**0 (2)**	0	–	15	2
	OR	2	0	4	0	**1 (1)**	–	0	0
	UT	–	–	–	–	–	**–**	–	–
	WV	35	4	14	17	0	–	**0 (2)**	1
	WY	8	0	3	3	0	–	2	**0 (1)**
Total	**CT**	**4 (6)**	14	61	50	18	–	82	13
	HI	15	**6 (6)**	7	8	5	–	16	1
	MSSA	62	8	**0 (3)**	25	16	–	39	5
	NH	54	8	28	**1 (5)**	5	–	37	4
	OR	19	5	17	5	**2 (2)**	–	5	1
	UT	–	–	–	–	–	**–**	–	–
	WV	84	16	41	42	5	–	**0 (6)**	3
	WY	14	1	5	5	1	–	4	**0 (1)**

[a]MSSA = Rhode Island and Vermont's Multi-State Science Assessment

Table 3. Number of Common Middle School Field-Test Items Administered and Calibrated, Spring 2018

	State	CT	HI	MSSA[a]	NH	OR	UT	WV	WY
Cluster	**CT**	**2 (6)**	12	22	26	19	44	77	5
	HI	11	**1 (0)**	3	6	6	0	9	1
	MSSA	23	3	**0 (1)**	9	1	7	22	2
	NH	26	6	10	**1 (2)**	7	0	17	3
	OR	19	6	1	7	**2 (2)**	0	5	1
	UT	48	0	7	0	0	**48 (52)**	43	0
	WV	83	10	21	18	6	48	**1 (9)**	2
	WY	5	1	2	3	1	0	2	**0 (0)**
Stand-Alone	**CT**	**2 (3)**	6	27	25	3	0	33	12
	HI	6	**8 (8)**	2	0	0	0	2	0
	MSSA	27	2	**0 (0)**	18	3	0	20	2
	NH	25	0	18	**0 (0)**	0	0	21	3
	OR	3	0	3	0	**0 (0)**	0	0	0
	UT	0	0	0	0	0	**0 (0)**	0	0
	WV	33	2	20	21	0	0	**0 (0)**	2
	WY	12	0	2	3	0	0	2	**0 (0)**
Total	**CT**	**4 (9)**	18	49	51	22	44	110	17
	HI	17	**9 (8)**	5	6	6	0	11	1
	MSSA	50	5	**0 (1)**	27	4	7	42	4
	NH	51	6	28	**1 (2)**	7	0	38	6
	OR	22	6	4	7	**2 (2)**	0	5	1
	UT	48	0	7	0	0	**48 (52)**	43	0
	WV	116	12	41	39	6	48	**1 (9)**	4
	WY	17	1	4	6	1	0	4	**0 (0)**

Note. [a]MSSA = Rhode Island and Vermont's Multi-State Science Assessment

Following the (operational) field-test administration, items underwent rubric validation and item data review, as described in Volume 2, Section 2.7.1, Rubric Validation, and Section 2.7.2, Data Review.

Table 4 presents the number of field-test items administered, the number of items rejected before or during rubric validation, the number of items submitted for data review, and the number of items rejected during data review.

Table 4. Number of Field-Test Item Administration, Rubric Validation, and Item Data Review, Spring 2018

Grade Band and Item Type	Number of Field-Test Items Administered	Number of Items Rejected Before/During Rubric Validation	Number of Items Submitted for Data Review	Number of Items Rejected at Data Review[a]	Number of Items Remaining
Elementary School	**153**	**3**	**65**	**13**	**137**
Cluster	86	3	24	5	78
Stand-Alone	67	0	41	8	59
Middle School	**241**	**16**	**102**	**24**	**201**
Cluster	171	12	65	15	144
Stand-Alone	70	4	37	9	57
Total	**394**	**19**	**167**	**37**	**338**

Note. [a]Figures in this column include three middle school clusters rejected after item data review

Table 5 summarizes the operational Shared Science Assessment Item Bank for each of the three science disciplines after adding the 2018 field-test items that passed rubric validation and item data review.

Table 5. Summary of Shared Science Assessment Item Bank, Spring 2018

Grade Band and Item Type	Science Discipline			Total[a]
	Earth and Space Sciences	*Life Sciences*	*Physical Sciences*	
Elementary School	**41**	**47**	**49**	**137**
Cluster	23	29	26	78
Stand-Alone	18	18	23	59
Middle School	**56**	**72**	**70**	**198**
Cluster	41	49	51	141
Stand-Alone	15	23	19	57
Total	**97**	**119**	**119**	**335**

[a]Total excludes three Utah-owned middle school clusters that aligned to Utah-specific standards

2. 2019 Field Tests

In 2019, a second wave of items was field tested in nine states. For three states (Hawaii, Idaho [elementary school only], and Wyoming), unscored field-test items were added as a separate segment to the operational scored legacy science test. An independent field test, in which students were administered a full set of items, was conducted for a sample of Idaho middle schools. In the remaining six states (Connecticut, New Hampshire, Oregon, Rhode Island, Vermont, and West Virginia), field-test items were administered as unscored items embedded among the operational items. In total, 88 item clusters and 156 stand-alone items were administered as field-test items in the elementary and middle school grade bands. Table 6 presents the number of field-test item clusters and stand-alone items administered in each grade band for each state.

Table 6. Number of Field-Test Items Administered, Spring 2019

Grade Band and Item Type	CT	HI	ID	MSSA[a]	NH	OR	WV	WY	Total
Elementary School	**47**	**31**	**53**	**42**	**18**	**27**	**18**	**16**	**117**
Cluster	18	19	20	17	0	16	10	5	50
Stand-Alone	29	12	33	25	18	11	8	11	67
Middle School	**56**	**23**	**53**	**46**	**28**	**26**	**26**	**15**	**127**
Cluster	14	9	17	10	4	9	8	5	38
Stand-Alone	42	14	36	36	24	17	18	10	89

Grade Band and Item Type	CT	HI	ID	MSSA[a]	NH	OR	WV	WY	Total
Total	**103**	**54**	**106**	**88**	**46**	**53**	**44**	**31**	**244**

[a]MSSA = Rhode Island and Vermont's Multi-State Science Assessment

For the three states with a separate field-test segment (i.e., states with a legacy science test), field-test forms were constructed using a balanced incomplete design and randomly assigned so that the group of students administered one form was comparable to the groups of students that were assigned other forms. For the independent field test, items were administered under a LOFT design, where the only blueprint constraint imposed was that students received four stand-alone items and two item clusters for each of the three science disciplines.

In the states with an operational test, field-test items were embedded within the test. Three states with an operational test—New Hampshire, Rhode Island, and Vermont—opted for a test in which operational items were grouped by science discipline. For these states, the field-test items were presented together in a fourth group of items. The sequence of the four sets of items (corresponding to the three science disciplines and a set of field-test items) was randomized across students. Three other states—Connecticut, Oregon, and West Virginia—opted for a test design in which the items were not grouped by discipline. In these three states, field-test items were administered at random positions throughout the test. A student received either a field-test item cluster or a set of five field-test stand-alone items.

A minimum sample size of 1,500 students per field-test item was targeted for any given state. Most items were administered in two or more states. Approximately 88.8% of the items met or exceeded the target sample size of 1,500 in at least one state, while 96.4% of the items had a sample size of at least 1,350 (10% of the target) in at least one state.

Table 7 and Table 8 present the number of item clusters and stand-alone items that were shared between the field-test pools of any two states. The numbers below the shaded cells represent the number of common field-test items between any two states, while the numbers above the shaded cells represent the number of common field-test items that survived rubric validation and were included in the calibration. In each of the shaded cells, the number outside the parentheses represents the number of unique field-test items administered only in the given state, and the number provided in parentheses represents the number of unique and/or common items that were calibrated with the data from that state only.

Table 7 presents the results for elementary schools and Table 8 presents the results for middle schools. The numbers of field-test items administered are slightly different from the numbers of field-test items at calibration because some items were rejected during rubric validation.

Table 7. Number of Common Elementary School Field-Test Items Administered and Calibrated, Spring 2019

	State	CT	HI	ID	MSSA[a]	NH	OR	WV	WY
Cluster	CT	**2 (2)**	2	10	3	0	2	1	4
	HI	2	**0 (0)**	3	8	0	14	2	0
	ID	10	3	**4 (4)**	0	0	1	3	3
	MSSA	3	8	0	**3 (3)**	0	9	4	1
	NH	0	0	0	0	**0 (0)**	0	0	0
	OR	2	14	1	9	0	**1 (1)**	0	0
	WV	1	2	3	4	0	0	**1 (0)**	1
	WY	4	0	3	1	0	0	1	**0 (0)**
Stand-Alone	CT	**5 (5)**	1	13	1	9	0	0	2
	HI	1	**0 (0)**	10	6	0	6	0	0
	ID	13	11	**1 (1)**	12	1	9	2	4
	MSSA	1	7	13	**3 (3)**	5	8	5	6
	NH	9	0	1	5	**2 (3)**	0	0	6
	OR	0	7	10	9	0	**1 (1)**	0	0
	WV	0	0	2	5	0	0	**1 (1)**	0
	WY	2	0	4	6	7	0	0	**0 (0)**
Total	CT	**7 (7)**	3	23	4	9	2	1	6
	HI	3	**0 (0)**	13	14	0	20	2	0
	ID	23	14	**5 (5)**	12	1	10	5	7
	MSSA	4	15	13	**6 (6)**	5	17	9	7
	NH	9	0	1	5	**2 (3)**	0	0	6
	OR	2	21	11	18	0	**2 (2)**	0	0
	WV	1	2	5	9	0	0	**2 (1)**	1
	WY	6	0	7	7	7	0	1	**0 (0)**

[a]MSSA = Rhode Island and Vermont's Multi-State Science Assessment

Table 8. Number of Common Middle School Field-Test Items Administered and Calibrated, Spring 2019

	State	CT	HI	ID	MSSA[a]	NH	OR	WV	WY
Cluster	CT	**5 (5)**	3	4	2	0	2	1	0
	HI	3	**0 (0)**	4	4	0	5	1	0
	ID	4	4	**2 (2)**	4	0	4	3	3
	MSSA	2	4	4	**1 (1)**	0	2	3	1
	NH	0	0	1	0	**3 (0)**	0	0	0
	OR	2	5	4	2	0	**1 (1)**	1	2
	WV	1	1	3	3	0	1	**0 (0)**	2
	WY	0	0	3	1	0	2	2	**0 (0)**
Stand-Alone	CT	**10 (9)**	2	13	9	10	3	6	0
	HI	2	**0 (0)**	9	9	0	6	3	0
	ID	13	9	**2 (2)**	11	1	12	6	5
	MSSA	9	9	11	**1 (1)**	6	11	9	7
	NH	10	0	2	6	**3 (1)**	0	0	2
	OR	3	6	12	11	0	**0 (0)**	2	7
	WV	6	3	6	9	1	2	**0 (0)**	0
	WY	0	0	5	7	2	7	0	**0 (0)**
Total	CT	**15 (14)**	5	17	11	10	5	7	0
	HI	5	**0 (0)**	13	13	0	11	4	0
	ID	17	13	**4 (4)**	15	1	16	9	8
	MSSA	11	13	15	**2 (2)**	6	13	12	8
	NH	10	0	3	6	**6 (1)**	0	0	2
	OR	5	11	16	13	0	**1 (1)**	3	9
	WV	7	4	9	12	1	3	**0 (0)**	2
	WY	0	0	8	8	2	9	2	**0 (0)**

[a]MSSA = Rhode Island and Vermont's Multi-State Science Assessment

Following the (operational) administration, field-test items underwent rubric validation and item data review, as described in Volume 2, Section, 2.7.1, Rubric Validation, and Section 2.7.2, Data Review.

Table 9 presents the number of field-test items administered, the number of items rejected before or during rubric validation, the number of items sent for data review, and the number of items rejected during data review.

Table 9. Number of Field-Test Item Administration, Rubric Validation, and Item Data Review, Spring 2019

Grade Band and Item Type	Number of Items Field Tested	Number of Items Rejected Before/During Rubric Validation	Number of Items Submitted for Data Review	Number of Items Rejected at Data Review	Number of Items Remaining[a]
Elementary School	**117**	**2**	**72**	**24**	**91**
Cluster	50	1	16	10	39
Stand-Alone	67	1	56	14	52
Middle School	**127**	**6**	**66**	**21**	**97**
Cluster	38	1	12	5	29
Stand-Alone	89	5	54	16	68
Total	**244**	**8**	**138**	**45**	**188**

Note. [a]Number of items remaining excludes five AI-scored items (four ICCR and one MSSA-owned) field tested in spring 2019 that were not brought to item data review.

Table 10 summarizes the Shared Science Assessment Item Bank after adding the field-test items that were administered in 2019 and passed rubric validation and item data review.

Table 10. Summary of Shared Science Assessment Item Bank, Spring 2019

Grade Band and Item Type	Science Discipline			Total[a]
	Earth and Space Sciences	*Life Sciences*	*Physical Sciences*	
Elementary School	**68**	**77**	**80**	**225**
Cluster	33	40	40	113
Stand-Alone	35	37	40	112
Middle School	**83**	**108**	**92**	**287**
Cluster	44	62	53	163
Stand-Alone	39	46	39	124
Total	**151**	**185**	**172**	**508**

3. 2021 Field Tests

In 2021, a third wave of items was field tested in 12 states. For one state (Wyoming), unscored field-test items were added as a separate segment to the operational scored legacy science test. An independent field test, in which students were administered a full set of items, was conducted in Idaho and Montana. In the remaining nine states (Connecticut, Hawaii, New Hampshire, North Dakota, Rhode Island, South Dakota, Utah, Vermont, and West Virginia), field-test items were administered as unscored items embedded among the operational items. In total, 166 item clusters and 207 stand-alone items were administered as field-test items in the elementary and middle school grade bands. Table 11 presents the number of field-test item clusters and stand-alone items administered in each grade band for each state. The numbers in parentheses in the column representing Montana indicates the field-test items owned by Montana.

Table 11. Number of Field-Test Items Administered, Spring 2021

Grade Band and Item Type	CT	HI	ID	MSSA[a]	MT	ND	NH	SD	UT	WV	WY	Total
Elementary School	**36**	**22**	**140**	**55**	**21 (21)**	**11**	**19**	**8**	**54**	**19**	**17**	**214**
Cluster	16	6	58	18	7 (7)	3	3	3	54	7	5	106
Stand-Alone	20	16	82	37	14 (14)	8	16	5	0	12	12	108
Middle School	**33**	**19**	**129**	**54**	**20 (18)**	**11**	**18**	**11**	**45**	**19**	**20**	**159**
Cluster	17	6	44	18	7 (6)	3	2	2	45	7	4	60
Stand-Alone	16	13	85	36	13 (12)	8	16	9	0	12	16	99
Total	**69**	**41**	**269**	**109**	**41 (39)**	**22**	**37**	**19**	**99**	**38**	**37**	**373**

Note. The numbers in parentheses indicate Montana-owned items.
[a]MSSA = Rhode Island and Vermont's Multi-State Science Assessment

For the state with a separate field-test segment (i.e., Wyoming), field-test forms were constructed using a balanced incomplete design and randomly assigned so that the group of students administered one form was comparable to the groups of students that were assigned other forms. For the states with independent field tests (i.e., Idaho and Montana), items were administered under a LOFT design, where the only blueprint constraint imposed was that students received four stand-alone items and two item clusters for each of the three science disciplines.

For the states with an operational test, field-test items were embedded within the test. Three states with an operational test—New Hampshire, Rhode Island, and Vermont—chose to administer a test in which operational items were grouped by science discipline. For these states, the field-test items were presented together as a fourth group of items. The sequence of the four sets of items (corresponding to the three disciplines and a set of field-test items) was randomized across students. Six other states—Connecticut, Hawaii, North Dakota, South Dakota, Utah, and West Virginia—opted for a test design in which the items were not grouped by discipline. In these states, field-test items were administered at random positions throughout the test. A student received either a field-test item cluster or a set of four field-test stand-alone items.

A minimum sample size of 1,500 students per field-test item was targeted for any given state. Most items were administered in two or more states. Approximately 96.7% of the items met or exceeded the target sample size of 1,500 in at least one state, while 99.1% of the items had a sample size of at least 1,350 (10% of the target) in at least one state.

Table 12 and Table 13 present the number of item clusters and stand-alone items that were shared between the field-test pools of any two states. The numbers below the shaded cells represent the number of common field-test items between any two states, and the numbers above the shaded diagonal elements represent the number of common field-test items that survived rubric validation and were included in the calibration. In each of the shaded diagonal elements, the number outside the parentheses represents the number of unique field-test items administered only in the given state, and the number in parentheses represents the number of unique and/or common items that were calibrated with the data from that state only. Table 12 presents the results for elementary schools, and Table 13 presents the results for middle schools. The numbers of field-test items administered were slightly different from the numbers of field-test items at calibration because some items were rejected during rubric validation.

Table 12. Number of Common Elementary School Field-Test Items Administered and Calibrated, Spring 2021

	State	CT	HI	ID	MSSA[a]	MT	ND	NH	SD	UT	WV	WY
Item Clusters	CT	3 (3)	0	13	0	0	0	0	0	0	0	0
	HI	0	1 (1)	3	0	0	0	0	0	0	1	0
	ID	13	4	3 (2)	5	5	2	0	2	20	1	4
	MSSA	0	0	6	2 (2)	2	0	0	0	7	0	0
	MT	0	0	5	2	0 (0)	0	0	0	0	0	0
	ND	0	0	2	0	0	0 (0)	0	1	0	1	0
	NH	0	0	0	0	0	0	0 (0)	0	0	3	0
	SD	0	0	2	0	0	1	0	0 (0)	0	2	0
	UT	0	0	20	8	0	0	0	0	25 (24)	0	2
	WV	0	1	1	0	0	1	3	2	0	1 (1)	0
	WY	0	0	4	0	0	0	0	0	2	0	0 (0)
Stand-Alone Items	CT	3 (3)	0	14	2	0	0	0	0	0	0	1
	HI	0	0 (0)	12	1	0	0	2	3	0	1	0
	ID	14	12	3 (3)	30	13	4	3	3	0	4	9
	MSSA	2	1	30	0 (0)	12	0	3	1	0	0	0
	MT	0	0	13	12	0 (0)	0	0	0	0	0	0
	ND	0	0	4	0	0	0 (0)	2	0	0	0	1
	NH	0	2	4	3	0	2	0 (0)	2	0	3	1
	SD	0	3	3	1	0	0	2	0 (0)	0	0	0
	UT	0	0	0	0	0	0	0	0	0 (0)	0	0
	WV	0	1	4	0	0	1	3	0	0	3 (3)	0
	WY	1	0	9	0	0	1	1	0	0	0	0 (0)
Total	CT	6 (6)	0	27	2	0	0	0	0	0	0	1
	HI	0	1 (1)	15	1	0	0	2	3	0	2	0
	ID	27	16	6 (5)	35	18	6	3	5	20	5	13

	State	CT	HI	ID	MSSA[a]	MT	ND	NH	SD	UT	WV	WY
	MSSA	2	1	36	2 (2)	14	0	3	1	7	0	0
	MT	0	0	18	14	0 (0)	0	0	0	0	0	0
	ND	0	0	6	0	0	0 (0)	2	1	0	1	1
	NH	0	2	4	3	0	2	0 (0)	2	0	6	1
	SD	0	3	5	1	0	1	2	0 (0)	0	2	0
	UT	0	0	20	8	0	0	0	0	25 (24)	0	2
	WV	0	2	5	0	0	2	6	2	0	4 (4)	0
	WY	1	0	13	0	0	1	1	0	2	0	0 (0)

[a]MSSA = Rhode Island and Vermont's Multi-State Science Assessment

Table 13. Number of Common Middle School Field-Test Items Administered and Calibrated, Spring 2021

	State	CT	HI	ID	MSSA[a]	MT	ND	NH	SD	UT	WV	WY
Item Clusters	CT	0 (0)	0	9	2	0	0	0	0	10	0	0
	HI	0	0 (0)	2	3	0	0	0	0	3	1	0
	ID	11	2	1 (1)	10	6	2	1	1	31	0	4
	MSSA	4	3	11	0 (0)	0	2	0	0	9	1	1
	MT	0	0	6	0	1 (1)	0	1	1	4	0	0
	ND	0	0	3	2	0	0 (0)	0	0	2	0	0
	NH	0	0	1	0	1	0	0 (0)	1	0	1	0
	SD	0	0	1	0	1	0	1	0 (0)	0	0	0
	UT	14	3	36	11	4	3	0	1	0 (0)	2	2
	WV	0	1	1	1	0	0	1	1	5	0 (0)	0
	WY	0	0	4	1	0	0	0	0	2	0	0 (0)
Stand-Alone Items	CT	2 (2)	0	12	2	0	0	0	3	0	0	2
	HI	0	0 (0)	10	1	0	0	0	0	0	2	0
	ID	13	10	2 (2)	29	10	6	12	7	0	5	15
	MSSA	2	1	29	0 (0)	10	2	1	1	0	2	4
	MT	0	0	12	10	0 (0)	0	0	0	0	0	0
	ND	0	0	7	2	0	0 (0)	1	0	0	0	0
	NH	0	0	12	1	0	1	0 (0)	2	0	1	3
	SD	3	0	7	1	0	0	2	0 (0)	0	3	4
	UT	0	0	0	0	0	0	0	0	0 (0)	0	0
	WV	0	2	6	3	0	1	1	3	0	0 (0)	0
	WY	2	0	15	4	0	0	3	4	0	0	0 (0)
Total	CT	2 (2)	0	21	4	0	0	0	3	10	0	2
	HI	0	0 (0)	12	4	0	0	0	0	3	3	0
	ID	24	12	3 (3)	39	16	8	13	8	31	5	19

	State	CT	HI	ID	MSSA[a]	MT	ND	NH	SD	UT	WV	WY
	MSSA	6	4	40	0 (0)	10	4	1	1	9	3	5
	MT	0	0	18	10	1 (1)	0	1	1	4	0	0
	ND	0	0	10	4	0	0 (0)	1	0	2	0	0
	NH	0	0	13	1	1	1	0 (0)	3	0	2	3
	SD	3	0	8	1	1	0	3	0 (0)	0	3	4
	UT	14	3	36	11	4	3	0	1	0 (0)	2	2
	WV	0	3	7	4	0	1	2	4	5	0 (0)	0
	WY	2	0	19	5	0	0	3	4	2	0	0 (0)

[a]MSSA = Rhode Island and Vermont's Multi-State Science Assessment

Following the administration, field-test items went through rubric validation and item data review, as described in Volume 2, Section, 2.7.1, Rubric Validation, and Section 2.7.2, Data Review.

Table 14 presents the number of field-test items administered in Montana, or another state, the number of items rejected before or during rubric validation, the number of items sent for data review, and the number of items rejected during data review. The numbers in parentheses indicate the field-test items owned by Montana.

Table 14. Number of Field-Test Item Administration, Rubric Validation, and Item Data Review, Spring 2021

Grade Band and Item Type	Number of Field-Test Items Administered	Number of Items Rejected Before/During Rubric Validation	Number of Items Submitted for Data Review	Number of Items Rejected at Data Review	Number of Items Remaining[a]
Elementary School	**214 (21)**	**7 (0)**	**100 (21)**	**19 (2)**	**188 (19)**
Cluster	106 (7)	5 (0)	24 (7)	7 (1)	94 (6)
Stand-Alone	108 (14)	2 (0)	76 (14)	12 (1)	94 (13)
Middle School	**159 (18)**	**15 (2)**	**87 (16)**	**13 (2)**	**129 (14)**
Cluster	60 (6)	10 (0)	22 (6)	5 (1)	43 (5)
Stand-Alone	99 (12)	5 (2)	65 (10)	8 (1)	86 (9)
Total	**373 (39)**	**22 (2)**	**187 (37)**	**32 (4)**	**317 (33)**

Note. The numbers in parentheses indicate Montana-owned items.

Table 15 summarizes the Shared Science Assessment Item Bank after addition of the field-test items that were administered in 2021 and passed rubric validation and item data review. The numbers in parentheses indicate the items owned by Montana.

Table 15. Summary of Shared Science Assessment Item Bank, Spring 2021

Grade Band and Item Type	Science Discipline			Total[a]
	Earth and Space Sciences	*Life Sciences*	*Physical Sciences*	
Elementary School	**136 (5)**	**128 (7)**	**149 (7)**	**413 (19)**
Cluster	65 (2)	66 (2)	76 (2)	207 (6)
Stand-Alone	71 (3)	62 (5)	73 (5)	206 (13)
Middle School	**114 (5)**	**156 (4)**	**137 (5)**	**407 (14)**
Cluster	55 (2)	76 (1)	67 (2)	198 (5)
Stand-Alone	59 (3)	80 (3)	70 (3)	209 (9)
Total	**250 (10)**	**284 (11)**	**286 (12)**	**820 (33)**

Note. The numbers in parentheses indicate Montana-owned items.

4. 2022 Field Tests

In 2022, a fourth wave of items was field tested in 13 states and one U.S. territory. In all of these locations—Connecticut, Hawaii, Idaho, Montana, New Hampshire, North Dakota, Oregon, Rhode Island, South Dakota, the U.S. Virgin Islands, Utah, Vermont, West Virginia, and Wyoming—the field-test items were administered as unscored items embedded among the operational items. In total, 217 item clusters and 254 stand-alone items were administered as field-test items in the elementary, middle, and high school grade bands. Table 16 presents the number of field-test item clusters and stand-alone items administered in each grade band for each state or territory. The numbers in parentheses in the column representing Montana indicate the field-test items owned by Montana.

Table 16. Number of Field-Test Items Administered, Spring 2022

Grade Band and Item Type	CT	HI	ID	MSSA[a]	MT	ND	OR	NH	SD	VI	UT	WV	WY	Total
Elementary School	**34**	**28**	**22**	**66**	**12 (12)**	**12**	**17**	**41**	**10**	**1**	**62**	**19**	**10**	**170**
Cluster	22	8	11	22	4 (4)	4	5	15	4	1	62	11	2	88
Stand-Alone	12	20	11	44	8 (8)	8	12	26	6	0	-	8	8	82
Middle School	**40**	**30**	**35**	**64**	**12 (12)**	**12**	**17**	**39**	**10**	**1**	**76**	**33**	**10**	**190**
Cluster	20	10	7	21	4 (4)	4	5	16	4	1	76	5	2	88
Stand-Alone	20	20	28	43	8 (8)	8	12	23	6	0	-	28	8	102
Total	**74**	**58**	**57**	**130**	**24 (24)**	**24**	**34**	**80**	**20**	**2**	**138**	**52**	**20**	**360**

Note. The numbers in parentheses indicate Montana-owned items.
[a]MSSA = Rhode Island and Vermont's Multi-State Science Assessment

In the states where an operational test was administered in spring 2022, field-test items were embedded within the test. Three states with an operational test—New Hampshire, Rhode Island, and Vermont—chose to administer a test in which operational items were grouped by science discipline. For these states, the field-test items were presented together as a fourth group of items. The sequence of the four sets of items (corresponding to the three disciplines and a set of field-test items) was randomized across students. Ten other states and one U.S. territory—Connecticut, Hawaii, Idaho, Montana, North Dakota, Oregon, South Dakota, Utah, the U.S. Virgin Islands, West Virginia, and Wyoming—opted for a test design in which the items were not grouped by discipline. In these 10 states and one U.S. territory, field-test items were administered at random positions throughout the test. A student received either a field-test item cluster or a set of four field-test stand-alone items. The test design for the Montana Science Assessment is discussed in Section 3.3, Test Design.

A minimum sample size of 1,500 students per field-test item was targeted for any given state or territory. Most items were administered in two states or one state and one territory. Approximately 61.6% of the items met or exceeded the target sample size of 1,500 in at least one state, while 88.0% of the items had a sample size of at least 1,350 (10% of the target) in at least one state. In addition, 98.3% of the items had a sample size of at least 1,200 in at least one state.

Table 17 and Table 18 present the number of item clusters and stand-alone items that were shared between the field-test pools of any two states (or territory). The numbers below the shaded diagonal elements represent the number of common field-test items between any two states, and the numbers above the shaded cells represent the number of common field-test items that survived rubric validation and were included in the calibration. In each of the shaded cells, the number outside the parentheses represents the number of unique field-test items administered only in the given state or territory, and the number in parentheses represents the number of unique and/or common items that were calibrated with the data from that state only. Table 17 presents the results for elementary schools and Table 18 presents the results for middle schools. The numbers of field-test items administered differ slightly from the numbers of field-test items at calibration because some items were rejected during rubric validation.

Table 17. Number of Common Elementary School Field-Test Items Administered and Calibrated, Spring 2022

	State	CT	HI	ID	MSSA[a]	MT	ND	NH	OR	SD	USVI	UT	WV	WY
Item Clusters	**CT**	0 (0)	0	3	1	0	0	0	3	0	0	15	0	0
	HI	0	0(0)	0	0	0	0	0	5	0	0	2	0	0
	ID	3	0	0(0)	3	0	0	0	0	0	0	5	0	0
	MSSA	1	0	3	0(0)	0	0	0	5	1	0	12	0	0
	MT	0	0	0	0	0(0)	0	0	0	0	0	4	0	0
	ND	0	0	0	0	0	0(0)	4	0	0	0	0	0	0
	NH	0	0	0	0	0	4	0(0)	0	0	1	1	0	0
	OR	3	6	0	5	0	0	0	0(0)	0	0	1	0	0
	SD	0	0	0	1	0	0	0	0	0 (0)	0	3	0	0
	USVI	0	0	0	0	0	0	1	0	0	0(0)	1	0	0
	UT	15	2	5	12	4	0	1	1	3	1	6 (6)	11	2
	WV	0	0	0	0	0	0	0	0	0	0	11	0(0)	0
	WY	0	0	0	0	0	0	0	0	0	0	2	0	0 (0)
Stand-Alone Items	**CT**	0 (0)	2	0	4	4	0	0	0	2	0	0	0	0
	HI	2	0(0)	3	7	0	0	0	8	0	0	0	0	0
	ID	0	3	0(0)	1	1	4	0	2	0	0	0	0	0
	MSSA	4	7	1	0(0)	3	0	0	7	4	0	0	8	8
	MT	4	0	1	3	0(0)	0	0	0	0	0	0	0	0
	ND	0	0	4	0	0	0(0)	3	1	0	0	0	0	0
	NH	0	0	0	0	0	3	1 (0)	7	0	0	0	0	0
	OR	0	8	2	8	0	1	7	0(0)	0	0	0	0	0
	SD	2	0	0	4	0	0	0	0	0 (0)	0	0	0	0
	USVI	0	0	0	0	0	0	0	0	0	0(0)	0	0	0
	UT	0	0	0	0	0	0	0	0	0	0	0(0)	0	0
	WV	0	0	0	8	0	0	0	0	0	0	0	0(0)	0

	State	CT	HI	ID	MSSA[a]	MT	ND	NH	OR	SD	USVI	UT	WV	WY
	WY	0	0	0	8	0	0	0	0	0	0	0	0	0 (0)
Total	CT	0(0)	2	3	5	4	0	0	3	2	0	15	0	0
	HI	2	0(0)	3	7	0	0	0	13	0	0	2	0	0
	ID	3	3	0(0)	4	1	4	0	2	0	0	5	0	0
	MSSA	5	7	4	0(0)	3	0	0	12	5	0	12	8	8
	MT	4	0	1	3	0(0)	0	0	0	0	0	4	0	0
	ND	0	0	4	0	0	0(0)	7	1	0	0	0	0	0
	NH	0	0	0	0	0	7	1 (0)	7	0	1	1	0	0
	OR	3	14	2	13	0	1	7	0(0)	0	0	1	0	0
	SD	2	0	0	5	0	0	0	0	0 (0)	0	3	0	0
	USVI	0	0	0	0	0	0	1	0	0	0(0)	1	0	0
	UT	15	2	5	12	4	0	1	1	3	1	6 (6)	11	2
	WV	0	0	0	8	0	0	0	0	0	0	11	0(0)	0
	WY	0	0	0	8	0	0	0	0	0	0	2	0	0 (0)

[a]MSSA = Rhode Island and Vermont's Multi-State Science Assessment

Table 18. Number of Common Middle School Field-Test Items Administered and Calibrated, Spring 2022

	State	CT	HI	ID	MSSA[a]	MT	ND	NH	OR	SD	USVI	UT	WV	WY
Item Clusters	CT	0 (0)	1	1	0	0	0	0	1	0	0	17	0	0
	HI	1	0 (0)	0	1	0	0	0	1	0	0	5	0	0
	ID	1	0	0 (0)	0	0	0	0	0	1	0	5	0	0
	MSSA	0	1	0	0 (0)	0	0	0	2	0	0	18	0	0
	MT	0	0	0	0	0 (0)	0	0	0	0	0	4	0	0
	ND	0	0	0	0	0	0 (0)	0	0	0	0	4	0	0
	NH	0	0	0	0	0	0	0 (0)	3	0	0	2	0	0
	OR	1	2	0	2	0	0	3	0 (0)	0	0	8	0	0
	SD	0	0	1	0	0	0	0	0	0 (0)	1	2	0	0
	USVI	0	0	0	0	0	0	0	0	1	0 (0)	1	0	0
	UT	17	6	5	18	4	4	2	8	3	1	2 (2)	5	2
	WV	0	0	0	0	0	0	0	0	0	0	5	0 (0)	0
	WY	0	0	0	0	0	0	0	0	0	0	2	0	0 (0)
Stand-Alone Items	CT	0 (0)	0	0	12	0	0	0	4	1	0	0	3	0
	HI	0	0 (0)	8	5	0	0	0	6	0	0	0	1	0
	ID	0	8	0 (0)	5	8	0	0	3	4	0	0	0	0
	MSSA	12	5	5	0 (0)	0	0	0	4	0	0	0	9	8
	MT	0	0	8	0	0 (0)	0	0	0	0	0	0	0	0
	ND	0	0	0	0	0	0 (0)	0	0	0	0	0	8	0
	NH	0	0	0	0	0	0	0 (0)	6	0	0	0	5	0
	OR	4	6	3	4	0	0	6	0 (0)	0	0	0	0	0
	SD	1	0	4	0	0	0	0	0	0 (0)	0	0	1	0
	USVI	0	0	0	0	0	0	0	0	0	0 (0)	0	0	0
	UT	0	0	0	0	0	0	0	0	0	0	0 (0)	0	0
	WV	3	1	0	9	0	8	6	0	1	0	0	0 (0)	0

	State	CT	HI	ID	MSSA[a]	MT	ND	NH	OR	SD	USVI	UT	WV	WY
	WY	0	0	0	8	0	0	0	0	0	0	0	0	0 (0)
Total	CT	0 (0)	1	1	12	0	0	0	5	1	0	17	3	0
	HI	1	0 (0)	8	6	0	0	0	7	0	0	5	1	0
	ID	1	8	0 (0)	5	8	0	0	3	5	0	5	0	0
	MSSA	12	6	5	0 (0)	0	0	0	6	0	0	18	9	8
	MT	0	0	8	0	0 (0)	0	0	0	0	0	4	0	0
	ND	0	0	0	0	0	0 (0)	0	0	0	0	4	8	0
	NH	0	0	0	0	0	0	0 (0)	9	0	0	2	5	0
	OR	5	8	3	6	0	0	9	0 (0)	0	0	0	0	0
	SD	1	0	5	0	0	0	0	0	0 (0)	1	2	1	0
	USVI	0	0	0	0	0	0	0	0	1	0 (0)	1	0	0
	UT	17	6	5	18	4	4	2	0	3	1	2 (2)	5	2
	WV	3	1	0	9	0	8	6	0	1	0	5	0 (0)	0
	WY	0	0	0	8	0	0	0	0	0	0	2	0	0 (0)

[a]MSSA = Rhode Island and Vermont's Multi-State Science Assessment

Following administration, field-test items went through rubric validation and item data review, as described in Volume 2, Section, 2.7.1, Rubric Validation, and Section 2.7.2, Data Review.

Table 19 presents the number of field-test items administered in Montana or another state (or territory), the number of items rejected before or during rubric validation, the number of items submitted for data review, and the number of items rejected during data review. The numbers in parentheses indicate the field-test items owned by Montana.

Table 19. Number of Field-Test Item Administration, Rubric Validation, and Item Data Review, Spring 2022

Grade Band and Item Type	Number of Field-Test Items Administered	Number of Items Rejected Before/During Rubric Validation	Number of Items Submitted for Data Review	Number of Items Rejected at Data Review	Number of Items Remaining
Elementary School	**170 (12)**	**3 (0)**	**82 (9)**	**14 (1)**	**153 (11)**
Cluster	88 (4)	1 (0)	18 (3)	4 (1)	83 (3)
Stand-Alone	82 (8)	2 (0)	64 (6)	10 (0)	70 (8)
Middle School	**190 (12)**	**4 (0)**	**94 (8)**	**26 (2)**	**160 (10)**
Cluster	88 (4)	3 (0)	26 (2)	13 (1)	72 (3)
Stand-Alone	102 (8)	1 (0)	68 (6)	13 (1)	88 (7)
Total	**360 (24)**	**7 (0)**	**176 (17)**	**40 (3)**	**313 (21)**

Note. The numbers in parentheses indicate Montana-owned items.

Table 20 summarizes the Shared Science Assessment Item Bank after adding the field-test items that were administered in 2022 and passed rubric validation and item data review. The numbers in parentheses indicate the items owned by Montana.

Table 20. Summary of Shared Science Assessment Item Bank, Spring 2022

Grade Band and Item Type	Science Discipline			Total[a]
	Earth and Space Sciences	*Life Sciences*	*Physical Sciences*	
Elementary School	**180 (10)**	**162 (10)**	**214 (10)**	**556 (30)**
Cluster	96 (4)	82 (3)	111 (2)	289 (9)
Stand-Alone	84 (6)	80 (7)	103 (8)	267 (21)
Middle School	**150 (9)**	**220 (8)**	**187 (7)**	**557 (24)**
Cluster	70 (2)	110 (3)	90 (3)	270 (8)
Stand-Alone	80 (7)	110 (5)	97 (4)	287 (16)
Total	**330 (19)**	**382 (18)**	**401 (17)**	**1113 (54)**

Note. The numbers in parentheses indicate Montana-owned items.
[a]Count excludes eight MOU items that do not align to the NGSS

Appendix 1-C
Calibration of the Shared Science Assessment Item Bank

TABLE OF CONTENTS

List of Tables

The Shared Science Assessment Item Bank was first calibrated in 2018 after the 2018 science test administrations concluded, and it was recalibrated in 2019 following the 2019 test administrations. The calibration sequences are documented in this appendix, which also includes details on scale linking and the creation of the anchor scale in 2019. The calibration of field-test items in 2021 and beyond as well as the calibration software are addressed.

1. 2018 Calibration Sequence

Table 1 provides an overview of the groups per grade band for the 2018 calibration.

Table 1. Groups Per Grade Band for the Spring 2018 Core Calibration

GROUP	ELEMENTARY SCHOOL	MIDDLE SCHOOL	HIGH SCHOOL
CONNECTICUT	X	X	X
HAWAII	X	X	X
NEW HAMPSHIRE	X	X	X
RHODE ISLAND	X	X	X
UTAH GRADE 6		X	
UTAH GRADE 7		X	
UTAH GRADE 8		X	
VERMONT	X	X	X
WEST VIRGINIA	X	X	

Items were calibrated in three steps for two reasons. First, the rubric validations for some states took place at a later date, and the student responses for the items owned by those states could not be included in the first round of calibrations without jeopardizing the reporting schedule of the two states with operational field tests (i.e., those two states did not have any of the items with late rubric validation in their item pool). Second, to divide the large set of items and assertions into more manageable portions, a separate calibration was conducted for two states with many items administered in those states only. Specifically, the following sequence of calibrations was conducted:

1. **Core Calibration.** The core calibration was performed on the following items:
    a. All item responses for New Hampshire and West Virginia. These states administered items from the following sources (as described in the state-sharing matrix in Table 2):
        i. ICCR item bank
        ii. Connecticut
        iii. Hawaii
        iv. Rhode Island

v. Vermont

vi. Utah

vii. West Virginia

A more detailed overlap of the common items at the time of the 2018 calibration was given in Section 1 of Appendix 1-B, Shared Science Assessment Item Bank: Field Testing (Table 2 -Table 4).

b. All item responses from Connecticut, Rhode Island, and Vermont excluding responses to Wyoming and Oregon items. These states administered items from the following sources:

i. ICCR item bank

ii. Connecticut

iii. Hawaii

iv. Rhode Island

v. Vermont

vi. Utah

vii. West Virginia

viii. Wyoming (items were treated as “not administered”; responses were replaced by missing code)

ix. Oregon (items were treated as “not administered”; responses were replaced by missing code)

c. Item responses from Hawaii to items also administered in another state (Hawaii items were used in Hawaii, Connecticut, Rhode Island, Vermont, and West Virginia).

d. Item responses from Utah to items also administered in another state (Utah items were used in Utah, Connecticut, Rhode Island, Vermont, and West Virginia) underwent core calibration. Utah tested middle school students only. One-third of students were selected at random to balance the large population size for Utah.

Table 2. Spring 2018 State-Sharing Matrix

Source Bank	CT	HI	MSSA[a]	NH	OR	UT	WV	WY
ICCR	X	X	X	X	X		X	X
Connecticut	X		X				X	
Hawaii	X	X	X				X	
MSSA	X		X				X	
Oregon	X		X		X			
Utah	X		X			X	X	
West Virginia	X		X				X	
Wyoming	X		X					X

Note. The core calibration provided parameters for all items used in New Hampshire and West Virginia.
[a]MSSA = Rhode Island and Vermont's Multi-State Science Assessment

2. **Calibration of State-Specific Items.** In terms of the calibration of state-specific items, both Hawaii and Utah had a substantial proportion of items that were administered only in Hawaii and Utah, respectively. Hawaii had both Hawaii and ICCR items in common with the states involved in the core calibration (Hawaii administered Hawaii and ICCR items only); whereas Utah had only Utah items in common (Utah administered Utah items only). The parameters for the unique Hawaii items depended on responses from Hawaii students only, and the parameters for the unique Utah items depended on responses from Utah students only. For both states, the state-specific items were calibrated through a separate calibration based on the state data only, with the items in common with the core states mentioned in Step 1 anchored to the estimates from Step 1. These calibrations were performed separately for each group under a single-group item response theory (IRT) model. The mean and variance of the groups were fixed to the estimated mean and variance from the core calibration.

3. **Calibration of States with Late Rubric Validation.** Oregon and Wyoming items were administered in some of the states involved in the core calibration (Connecticut, Rhode Island, and Vermont) but could not be calibrated in Step 1 because of their late rubric validation dates. In a later stage, items from Oregon and Wyoming were calibrated by:

    a. adding Oregon and Wyoming student responses to the core calibration;

    b. keeping the responses from Connecticut, Rhode Island, and Vermont to Wyoming and Oregon items (as opposed to treating them as missing in Step 1);

    c. removing the responses from Hawaii, Utah, New Hampshire, and West Virginia, who did not administer Oregon or Wyoming items (as the item parameters for the Oregon and Wyoming items did not depend on the students from these states); and

    d. fixing the parameters of all other items to the values obtained in Step 1 and the group means and standard deviations that were estimated in Step 1.

2. 2019 CALIBRATION

Calibration was performed in two steps. First, CAI calibrated all items in operational use in 2019, for which 1,000 or more student responses were available (among these, there were 1,500 or more student responses for all but three items). In this step, only the data from states with an operational test were included. Table 3 provides an overview of the groups per grade band for this first calibration. All students who attempted the test were included in the calibration. The assertions of skipped items were scored as incorrect. Note that only Rhode Island allowed students to skip items. Out of a total of 438 items, there were nine items administered as operational items in 2019, for which the sample size was smaller than 1,000.

Table 4 and Table 5 present the number of operational item clusters and stand-alone items that were shared between the item pools of any two states. The numbers below the shaded cells represent the number of all the operational items administered, and the numbers above the shaded cells represent the number of common operational items at the time of the 2019 calibration. The shaded cells represent the number of operational items administered only in the given state (the number of unique operational items at the time of calibration are provided in parentheses). Since the items that were administered but not calibrated were administered only in one state, the numbers above the diagonal are the same as the numbers below the diagonal.

Table 4 presents the results for elementary schools, and Table 5 presents the results for middle schools. The numbers at the operational administration are slightly different from the numbers at the calibration because items with sample sizes smaller than 1,000 were excluded from the calibration.

Table 3. Groups Per Grade Band for the Spring 2019 Calibration of Operational Items

Group	Elementary School	Middle School	High School
Connecticut	X	X	X
New Hampshire	X	X	X
Oregon	X	X	X
Rhode Island	X	X	X
Vermont	X	X	X
West Virginia	X	X	

Table 4. Number of Common Elementary School Operational Items Administered and Calibrated, Spring 2019

	State	CT	MSSA[a]	NH	OR	WV
Cluster	CT	**1 (1)**	44	24	42	55
	MSSA	44	**0 (0)**	17	37	41
	NH	24	17	**0 (0)**	14	27
	OR	42	37	14	**0 (0)**	41
	WV	55	41	27	41	**1 (1)**
Stand-Alone	CT	**3 (3)**	34	26	30	47
	MSSA	34	**0 (0)**	20	23	32
	NH	26	20	**0 (0)**	14	25
	OR	30	23	14	**0 (0)**	25
	WV	47	32	25	25	**1 (1)**
Total	CT	**4 (4)**	78	50	72	102
	MSSA	78	**0 (0)**	37	60	73
	NH	50	37	**0 (0)**	28	52
	OR	72	60	28	**0 (0)**	66
	WV	102	73	52	66	**2 (2)**

[a]MSSA = Rhode Island and Vermont's Multi-State Science Assessment

Table 5. Number of Common Middle School Operational Items Administered and Calibrated, Spring 2019

	State	CT	MSSA[a]	NH	OR	WV
Cluster	CT	**3 (3)**	26	24	54	92
	MSSA	26	**0 (0)**	11	14	21
	NH	24	11	**1 (1)**	9	18
	OR	54	14	9	**2 (2)**	56
	WV	92	21	18	56	**12 (4)**
Stand-Alone	CT	**0 (0)**	42	26	34	50
	MSSA	42	**0 (0)**	25	30	37
	NH	26	25	**0 (0)**	16	21
	OR	34	30	16	**1 (0)**	29
	WV	50	37	21	29	**0 (0)**
Total	CT	**3 (3)**	68	50	88	142
	MSSA	68	**0 (0)**	36	44	58
	NH	50	36	**1 (1)**	25	39
	OR	88	44	25	**3 (2)**	85
	WV	142	58	39	85	**12 (4)**

[a]MSSA = Rhode Island and Vermont's Multi-State Science Assessment

In Step 2, the field-test items were calibrated. The calibration included the operational items that were calibrated in Step 1 and the field-test items across all states in which they were administered. All students who attempted at least one field-test item were included in the calibration. Table 6 provides an overview of the groups per grade band for calibration of the field-test items.

Table 6. Groups Per Grade Band for the Spring 2019 Calibration of Field-Test Items

Group	Elementary School	Middle School	High School
Connecticut	X	X	X
Hawaii	X	X	X
Idaho	X	X	
New Hampshire	X	X	X
Oregon	X	X	X
Rhode Island	X	X	X
Vermont	X	X	X
West Virginia	X	X	
Wyoming	X	X	X

3. LINKING THE 2018 SCALE TO THE 2019 SCALE

The item parameter estimates obtained from the 2018 student responses were highly correlated with the item parameters obtained from the 2019 student responses. For item difficulties, the correlation between the 2018 and 2019 estimates was 0.993 for elementary school, 0.986 for middle school, and 0.994 for high school. For the standard deviations of the clusters, these correlations were 0.971 for elementary school, 0.972 for middle school, and 0.964 for high school. These high correlations indicate that items functioned similarly in 2018 and 2019. Nevertheless, item parameters from separate calibrations cannot be directly compared because the scale of an IRT model was not determined. In the multigroup Rasch testlet model, the only scale indeterminacy was the origin of the scale. The models can be identified by setting the mean of the overall proficiency variable θ to zero for the reference distribution. As a result, the 2018 and 2019 variable θ and item parameters were on the same scale except for an overall shift parameter B. Specifically, the 2018 scale can be linked to the 2019 scale as follows:

$$P\left(z_{ij}|\theta_{j\ 2018}, u_{jg}; b_{i\ 2018}\right) = \frac{\exp\left(\theta_{j\ 2018} + u_{jg} - b_{i\ 2018}\right)}{1 + \exp\left(\theta_{j\ 2018} + u_{jg} - b_{i\ 2018}\right)}$$

$$= \frac{exp\left(\theta_{j\ 2018} + B + u_{jg} - b_{i\ 2018} - B\right)}{1 + exp\left(\theta_{j\ 2018} + B + u_{jg} - b_{i\ 2018} - B\right)}$$

$$= \frac{exp\left(\theta_{j\ 2019} + u_{jg} - b_{i\ 2019}\right)}{1 + exp\left(\theta_{j\ 2019} + u_{jg} - b_{i\ 2019}\right)}.$$

Because $\theta_{j\ 2019} = \theta_{j\ 2018} + B$, the population means of θ must be transformed accordingly,

$$\theta_{j\ 2019} \sim N\ (\mu_{k\ 2018} + B, \sigma_k^2) \text{ and}$$

$$\theta_{j\ 2018} \sim N\ (\mu_{k\ 2018}, \sigma_k^2).$$

Item parameters based on 2018 student responses were expressed on the 2019 scale by adding the constant B to the 2018 item parameter. The 2018 parameters were expressed on the 2019 scale for items that were part of the pool in both 2018 and 2019 but were not administered in any states in 2019 (13 items), and for items that were administered in 2019 but the number of student responses from the 2019 assessments was lower than 1,000 (nine items). Therefore, the linking process was performed for 22 items only.

All items that were operational in 2019 were also administered in 2018. Therefore, the shift parameter B was estimated from a separate calibration of the 2019 operational items using the 2019 student responses (from the six operational states), but with the item parameters fixed to the estimates obtained from the 2018 calibrations. By fixing a subset of the item parameters, the model is identified so that the means and variances of θ can be estimated for all groups. Parameter B can be obtained by equating the overall mean of θ across all groups for the 2019 student response data from the free calibration (i.e., the 2019 overall mean expressed on the 2019 scale) to the overall mean of θ across all groups for the 2019 student response data from the calibration with items anchored to their 2018 parameters values (i.e., the 2019 overall mean expressed on the 2018 scale):

$$\frac{1}{K}\sum_{k=1}^{K}\mu_{k\ 2019} = \frac{1}{K}\sum_{k=1}^{K}(\mu_{k\ 2018} + B).$$

Therefore, an estimate of parameter B can be obtained as

$$\hat{B} = \frac{1}{K}\sum_{k=1}^{K}(\hat{\mu}_{k\ 2019} - \hat{\mu}_{k\ 2018}).$$

Table 7 presents the estimated means of θ under both the free and anchored calibrations and the number of students per state. Table 7 also presents the overall means and estimated shift in parameter B. Note that the parameters for three items were not anchored, but instead were freely estimated together with the means and variances in the anchored calibration. The reason for not treating these items as common items across the 2018 and 2019 administrations is that they had an omit rate of 4% or higher for the last item interaction in the 2018 administration in at least one state. In 2019, these interactions could no longer be omitted because all interactions of an item needed to be responded to in states where skipping was not allowed (i.e., all states excluding Rhode Island). Therefore, out of an abundance of caution, these three items are not anchored to their 2018 parameter values.

Table 7. Estimated Latent Means and Number of Students Per State

Group	Elementary School			Middle School		
	$\hat{\mu}_{k\ 2019}$	$\hat{\mu}_{k\ 2018}$	*N*	$\hat{\mu}_{k\ 2019}$	$\hat{\mu}_{k\ 2018}$	*N*
Connecticut	0.0000	0.0518	38,549	0.0000	0.0234	39,347
New Hampshire	0.0631	0.1083	13,187	0.0940	0.1108	12,060
Oregon	-0.0101	0.0096	44,989	0.0028	0.0156	42,043
Rhode Island	-0.0312	0.0142	10,751	-0.1044	-0.0692	10,306
Vermont	0.1069	0.1504	6,017	0.0781	0.1133	5,894
West Virginia	-0.1970	-0.1529	19,540	-0.3012	-0.2783	19,043
	$\frac{1}{K}\sum_{k=1}^{K}\hat{\mu}_{k\ 2019}$	$\frac{1}{K}\sum_{k=1}^{K}\hat{\mu}_{k\ 2018}$	$\hat{B}$	$\frac{1}{K}\sum_{k=1}^{K}\hat{\mu}_{k\ 2019}$	$\frac{1}{K}\sum_{k=1}^{K}\hat{\mu}_{k\ 2018}$	$\hat{B}$
Overall	-0.0114	0.0303	-0.0416	-0.0385	-0.0141	-0.0244

4. Calibration of Field-Test Items in 2021 and Beyond

Starting in 2021, field-test items were calibrated with one multigroup calibration per grade band. In each calibration, the parameters of the operational items were fixed to their bank values (anchor items), and the item parameters of the field-test items, as well as the mean and variance of each group were estimated using the marginal maximum likelihood (MML) method. The calibration included the field-test items across all states in which the items were administered. All students who attempted at least one field-test item were included in the calibration. In 2021 and 2022, the same groups were included for each grade band for the field test calibration. Refer to Table 8 and

Table 9 for an overview of the groups per grade band for calibration of the field-test items in 2021 and 2022, respectively.

In 2021, all but 12 items were calibrated on at least 1,500 student responses, and all but one item had a sample size larger than 1,200. The item with fewer than 1,200 responses had a sample size of 981 and was an interim item. In 2022, all but 64 items were calibrated on at least 1,500 student responses, and all but nine items were calibrated on at least 1,200 responses. The nine items with fewer than 1,200 responses were either Oregon legacy items or interim items.

Table 8. Groups Per Grade Band for the Spring 2021 Calibration of Field-Test Items

Group	Elementary School	Middle School	High School
Connecticut	X	X	X
Hawaii	X	X	X
Idaho	X	X	X
Montana	X	X	
North Dakota	X	X	X
New Hampshire	X	X	X
Rhode Island	X	X	X
South Dakota	X	X	X
Utah	X	X	
Vermont	X	X	X
West Virginia	X	X	
Wyoming	X	X	X

Table 9. Groups Per Grade Band for the Spring 2022 Calibration of Field-Test Items

Group	Elementary School	Middle School	High School
Connecticut	X	X	X
Hawaii	X	X	X
Idaho	X	X	X
Montana	X	X	
North Dakota	X	X	X
New Hampshire	X	X	X
Oregon	X	X	X
Rhode Island	X	X	X
South Dakota	X	X	X
Utah	X	X	
Vermont	X	X	X
West Virginia	X	X	
Wyoming	X	X	X

5. Calibration Software

In 2018 and 2019, the IRT models were fitted using the Bayesian networks with the logistic regression (BNL) suite of Matlab functions (Rijmen, 2006) and flexMIRT (Cai, 2017). The resulting parameters from BNL were used as starting values for flexMIRT to reduce the estimation time for flexMIRT. The flexMIRT estimates were taken to be the operational parameters, except for the middle school items calibrated in 2018 during the core calibration. For the 2018 core calibration of middle-school items, flexMIRT did not converge after several weeks, and the estimates obtained from BNL were used as operational parameters. Note that the parameters estimates were very similar across software packages.

Starting in 2021, Cambium Assessment IRT (CAIRT) was used for all calibrations because the estimation time in flexMIRT became prohibitive. CAIRT was specifically developed by CAI to calibrate the multigroup Rasch model on very large data sets. It relies on the same estimation methods as BNL. CAI has cross-validated parameter estimates from CAIRT with BNL and flexMIRT under various scenarios (Rijmen, Liao, & Lin, 2021). In 2023, field-test items were calibrated in CAIRT using the same procedure used in 2021.

6. References

Cai, L. (2017). flexMIRT®: Flexible multilevel multidimensional item analysis and test scoring (version 3.51) [computer software]. Chapel Hill, NC: Vector Psychometric Group.

Rijmen, F. (2006). *BNL: A Matlab toolbox for Bayesian networks with logistic regression nodes* (Technical Report). Amsterdam: VU University Medical Center.

Rijmen, F., Liao, D., & Lin, Z. (2021). *The Rasch testlet model for the calibration of three-dimensional science assessments: A software comparison* [White paper]. Washington, DC: Cambium Assessment, Inc.

Appendix 1-D

Distribution of Scale Scores and Performance Levels

Distribution of Scale Scores and Performance Levels

Table D-1. Scale Score Mean and Standard Deviation by Grade, Science

	Grade	
	5	*8*
Number of Students	11,068	11,159
Mean Scale Score	498.66	800.32
SD of Scale Score	27.38	27.69

Table D-2. Proportion of Students in Each Performance Level by Grade, Science

Performance Level	**Grade**	
	5	*8*
Number of Students	11,068	11,159
Level 1	0.22	0.23
Level 2	0.37	0.36
Level 3	0.28	0.24
Level 4	0.13	0.17

Appendix 1-E

Distribution of Scale Scores by Science Discipline

Distribution of Scale Scores by Science Discipline

Table E-1. Science Disciplines

Grade	Discipline
5, 8, 11	Physical Sciences (PS) Life Sciences (LS) Earth & Space Sciences (ESS)

Table E-2. Overall Discipline Score Mean and Standard Deviation, Grade 5 Science

N	Scale Score	Discipline		
		Physical Sciences	*Life Sciences*	*Earth and Space Sciences*
11,068	Mean	498.33	499.62	497.43
	SD	30.61	31.75	33.98

Table E-3. Overall Discipline Score Mean and Standard Deviation, Grade 8 Science

N	Scale Score	Discipline		
		Physical Sciences	*Life Sciences*	*Earth and Space Sciences*
11,159	Mean	801.60	799.40	799.15
	SD	30.25	31.31	33.26

Appendix 1-F

Distribution of Scale Scores and Performance Levels by Subgroup

Distribution of Scale Scores and Performance Levels by Subgroup

Table F-1. Mean and Standard Deviation of Scale Scores by Subgroup

Group	Grade 5			Grade 8		
	N	*Mean*	*SD*	*N*	*Mean*	*SD*
All Students	11,068	498.66	27.38	11,159	800.32	27.69
Female	5,406	497.91	26.66	5,384	800.17	25.80
Male	5,662	499.38	28.03	5,775	800.45	29.34
African American	88	495.46	27.27	57	797.97	26.36
American Indian/Native Alaskan	1,089	476.89	23.94	1,096	779.55	25.04
Asian	80	506.71	25.47	77	802.82	28.53
Hawaiian/Pacific Islander	16	493.51	26.23	25	793.17	27.38
Hispanic	609	494.98	26.89	660	787.97	25.45
Multi-Racial	536	501.93	26.55	532	803.71	26.82
White	8,650	497.39	25.78	8,712	787.89	26.70
Limited English Proficiency	374	477.51	24.42	458	775.14	26.06
Special Education	1,452	477.19	25.67	1,292	777.59	25.36
Economically Disadvantaged	4,856	489.76	26.81	4,070	789.96	26.67

Table F-2. Percentage of Performance Level by Subgroup

Group	Grade 5					Grade 8				
	N	*L1*	*L2*	*L3*	*L4*	*N*	*L1*	*L2*	*L3*	*L4*
All Students	11,068	22%	37%	28%	13%	11,159	23%	36%	24%	17%
Female	5,406	22%	38%	28%	11%	5,384	21%	39%	25%	15%
Male	5,662	22%	35%	29%	14%	5,775	25%	34%	23%	18%
African American	88	26%	42%	22%	10%	57	40%	37%	19%	4%
American Indian/Native Alaskan	1,089	52%	35%	11%	1%	1,096	52%	35%	11%	3%
Asian	80	15%	30%	39%	16%	77	21%	31%	30%	18%
Hawaiian/Pacific Islander	16	19%	38%	31%	12%	25	40%	36%	12%	12%
Hispanic	609	26%	39%	27%	8%	660	32%	38%	19%	11%
Multi-Racial	536	24%	40%	25%	10%	532	24%	40%	23%	13%
White	8,650	18%	37%	31%	15%	8,712	19%	36%	26%	19%
Limited English Proficiency	374	52%	33%	13%	2%	458	60%	29%	8%	3%
Special Education	1,452	54%	33%	10%	3%	1,292	56%	32%	8%	4%
Economically Disadvantaged	4,856	33%	39%	21%	7%	4,070	36%	38%	17%	9%

Montana Science Assessment

2022–2023

Volume 2:
Test Development

OFFICE OF PUBLIC INSTRUCTION
ELSIE ARNTZEN, STATE SUPERINTENDENT

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

1. INTRODUCTION

Montana adopted the Montana Science Content Standards in 2016. The new standards are built on the three-dimensional framework described in *A Framework for K–12 Science Education* (National Research Council, 2012) and closely resemble the Next Generation Science Standards (NGSS). To measure the new standards, the Montana Office of Public Instruction (OPI) and its assessment vendor, Cambium Assessment, Inc. (CAI; formerly the American Institutes for Research [AIR]) developed and administered the Montana Science Assessment (MSA), which was administered operationally for the first time in 2021–2022. The MSA measures the science knowledge and skills of Montana students in grades 5 and 8 as an online assessment, constructed with an adaptive test design that makes use of several technology-enhanced item types.

Additional details on the implementation of the new assessments can be found in Volume 1, Annual Technical Report.

The interpretation, usage, and validity of test scores rely heavily on how the MSA was developed. This volume of the technical report provides details on the test development process that contributes to the validity of the test scores. Specifically, this volume provides evidence to support the following:

- The item specifications, which offered detailed guidance for item writers and reviewers in order to ensure that science items were aligned to the performance expectations (PEs) they were intended to measure,
- The item development procedures employed for the Montana Science Assessment were consistent with industry standards,
- The development and maintenance of the Shared Science Assessment Item Bank, which contains item clusters and stand-alone items that cover the range of measured standards, grade-level difficulties, and levels of cognitive engagement, and
- The test design summary/blueprint stipulated the range of operational items from each item type and content category required for each test administration. This document was implemented using the item-selection algorithm for science.

Note that for the science assessments, as outlined in Volume 1, Annual Technical Report, CAI collaborated with a group of states that share common item-development processes. In addition to developing items for each of those states, CAI developed and maintains the Independent College and Career Readiness (ICCR) item bank, which consists of items developed according to the same principles that are followed for developing the items owned by each of the collaborator states. This volume of the technical report focuses on the general test development activities.

For the MSA, items are drawn from the Shared Science Assessment Item Bank, which consists of ICCR items, items owned by Montana, and items owned by several other states that have signed a Memorandum of Understanding (MOU) to share content, leadership, and new ideas and methods. Specifically, all items developed under the MOU underwent the same development process. For the remainder of this volume, the term *item bank* will refer to all items developed under the MOU unless stated otherwise.

1.1 Claim Structure

The goals, uses, and claims that the Shared Science Assessment Item Bank and subsequent tests were designed to support were identified in a series of collaborative meetings held over August 22–23, 2016. The overarching goal of these meetings was to support the development of statewide summative assessments using science content that measures the three-dimensional science standards based on *A Framework for K–12 Science Education* (National Research Council, 2012).

To this end, CAI invited content and assessment leaders from 10 states, as well as four nationally recognized experts, Aneesha Badrinarayan, Rodger Bybee, Peter McLaren, and Brett Moulding, who helped author the NGSS. Two nationally recognized psychometricians, Laurie Wise, Ph.D. and Tom Hirsch, Ph.D., also participated.

CAI staff and the participating states collaborated to develop items and test specifications designed to measure the three-dimensional science standards. In general, the item specifications were accompanied by sample item clusters that met those specifications. At the time, some standards did not have sample item clusters available. All specifications and sample item clusters were reviewed by state content experts and committees of educators in at least one of the states.

1.2 Underlying Principles Guiding Development

The Shared Science Assessment Item Bank was established using a highly structured, evidence-centered design. The process began with detailed item specifications. The specifications, discussed in Section 2.2, Item Specifications, described the interaction types that could be used, gave guidelines for targeting the appropriate cognitive engagement, offered suggestions for controlling item difficulty, and provided sample items.

Items were written with the goal that virtually every item would be accessible to all students, either by itself or in conjunction with accessibility tools, such as text-to-speech (TTS), translations, or assistive technologies. This goal is supported by the delivery of the items on CAI's Test Delivery System (TDS), which has received Web Content Accessibility Guidelines (WCAG) 2.0 AA certification, offers a wide array of accessibility tools and is compatible with most assistive technologies.

Item development supported the goal of high-quality item clusters and stand-alone items through rigorous development processes managed and tracked by a content development platform, which ensured that all items flowed through the correct sequence of reviews and captured every comment and change to each item.

CAI sought to ensure that the items measured the PEs in a fair and meaningful way by engaging educators, state officials, content experts, and fairness, bias, and sensitivity experts at each step of the process. Educators evaluated the alignment of the items to the PEs and offered guidance and suggestions for improvement. These educators participated in reviewing items for fairness and sensitivity. Following item field-testing, educators engaged in rubric validation, a process that refines rule-based rubrics upon review of student responses.

Combined, these principles and the processes that support them have been incorporated into an item bank that measures the PEs with fidelity and does so in a way that minimizes construct-irrelevant variance and reduces barriers to access. These processes are detailed in the chapters that follow.

1.3 Organization of This Volume

This volume is organized into the following three sections:

1. An overview of the science item development process that supports the validity of the claims that science tests are designed to support.
2. An overview of the Shared Science Assessment Item Bank, the types of assessments the item bank is designed to support, and methods for refreshing the item bank.
3. A description of the test construction process for the Montana Science Assessment, including the blueprint, the test design, an evaluation of simulated test sessions, the operational blueprint match results, and the item exposure rates

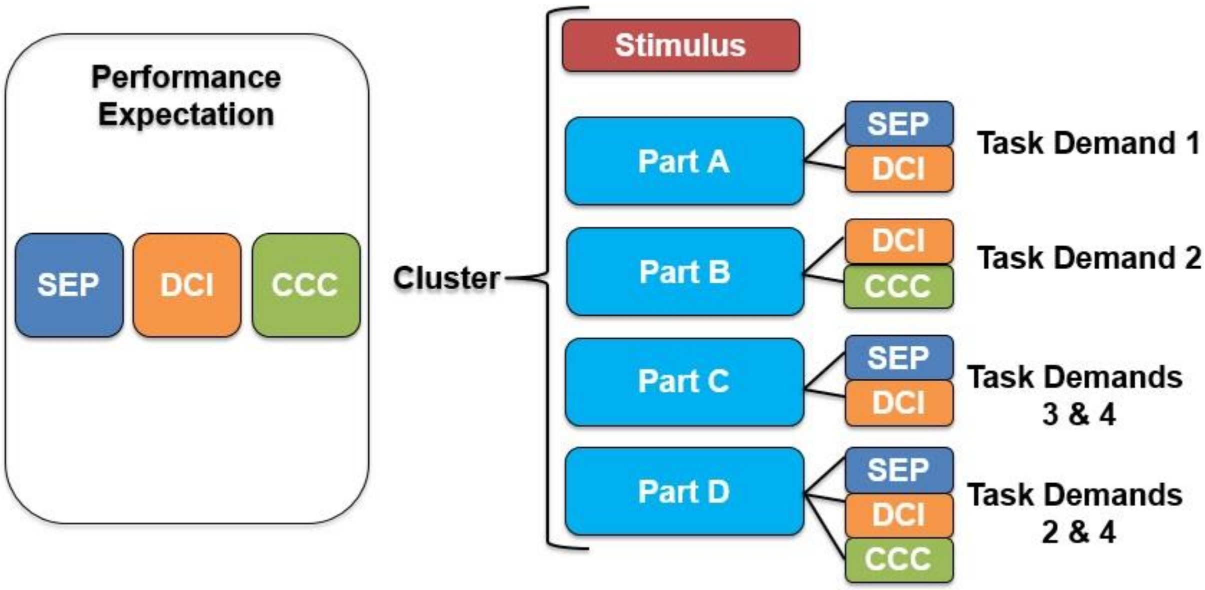
2. Item Development Process That Supports Validity of Claims

2.1 Overview

Cambium Assessment, Inc. (CAI) developed the Shared Science Assessment Item Bank in collaboration with the Memorandum of Understanding (MOU) participants by using a rigorous, structured process that engaged stakeholders at critical junctures.

A performance expectation is a point in a three-dimensional space formed by three dimensions of science learning: crosscutting concepts (CCCs), science and engineering practices (SEPs), and disciplinary core ideas (DCIs). That is, a performance expectation (PE) is characterized by a specific CCC, SEP, and DCI. When the MOU states first convened, many sessions were spent discussing how to assess these new three-dimensional standards. These group sessions are where the idea of an item cluster was conceived. An *item cluster* consists of a stimulus (scientific phenomenon) associated with multiple parts. Each of these parts contain questions that allow the student to explore the phenomenon. Each of the parts assess at least two dimensions, and the entire item cluster assesses a student on all three dimensions for a specific PE. Figure 1 is a visual representation of the structure of a three-dimensional cluster.

Figure 1. Structure of Three-Dimensional Item Clusters



Each part of an item cluster contains questions that require the student to interact with the item cluster. There are many different interactions that can be included in an item cluster. Section 3.1, Current Composition of the Shared Science Assessment Item Bank, describes and lists all of the different interactions available. The interactions used in an item cluster are chosen intentionally to best assess different aspects of the three-dimensional construct.

Figure 2 provides an example of an item cluster that has a phenomenon, five parts, and eight interactions; each part of an item cluster assesses multiple dimensions.

Figure 2. Example of an NGSS Item Cluster

A student rings a doorbell. When the person inside the house is on the main floor, he can easily hear the doorbell. When he is upstairs, though, he cannot so easily hear the doorbell.

Figure 1 shows the circuit of a simple doorbell when it is on (pressed) and off (not pressed).

Figure 1. Simple Doorbell Circuit

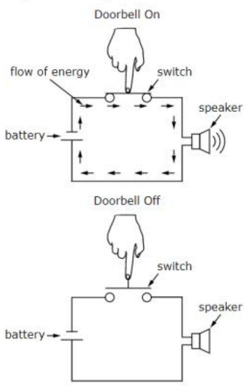


Table 1 shows the types of doorbell speakers available and their cost, in dollars ($).

Table 1. Types of Speakers and Cost

Speaker	Cost ($)
Bell	11
Buzzer	17
Chimes	25

Part A

Click on each blank box and select a phrase to describe what is happening to the energy at each part of the circuit when the doorbell is turned on.

Parts	Energy Pathway when Doorbell Is on
Battery	
Wires	
Speaker	

Part B

Use the simulation to select the materials necessary to conduct fair experiments and create a doorbell that can be heard fro upstairs and costs less than $40. The student can only hear a doorbell from upstairs if it is loud or very loud.

- Select the speaker, battery, and switch to determine the overall cost and loudness of the doorbell.
- Then click Run Trial.
- The cost of wire has already been included in the total cost.
- You must complete **two** trials.
- You may run up to **five** trials.
- Click the trash can icon if you want to delete a trial and generate new data.

Speaker Bell
Battery 12
Switch Rectangular
Run Trial

Trial	Speaker	Battery (V)	Switch	Loudness	Cost ($)

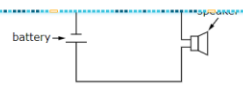


Table 1 shows the types of doorbell speakers available and their cost, in dollars ($).

Table 1. Types of Speakers and Cost

Speaker	Cost ($)
Bell	11
Buzzer	17
Chimes	25

Table 2 shows the types of batteries available based on their voltage (V), the amount of power each produces, and their cost.

Table 2. Types of Batteries and Cost

Battery (V)	Amount of Power	Cost ($)
12	A lot	27
9	Average	3
1.5	A little	1

Table 3 shows the types of switches and their cost.

Table 3. Types of Switches and Cost

Switches	Cost ($)
Rectangular	4
Circular	5
Lighted	11

Your Task

In the questions that follow, you will design a main-floor doorbell that can be heard from upstairs in a house.

Part C

Select **all** of the trials that meet the criteria for being heard upstairs and cost less than $40.

- ☐ Trial 1
- ☐ Trial 2
- ☐ Trial 3
- ☐ Trial 4
- ☐ Trial 5
- ☐ None

Part D

Click on the blank boxes and select words or phrases to predict what will happen to the loudness of the doorbell when the battery power increases.

The loudness of the doorbell will [ ] because [ ]

Part E

Select **two** trials that support the relationship between the loudness of the doorbell and the power of the battery.

- ☐ Trial 1
- ☐ Trial 2
- ☐ Trial 3
- ☐ Trial 4
- ☐ Trial 5
- ☐ Cannot be determined

This item cluster is aligned to the NGSS PE of 4-PS3-4: Apply scientific ideas to design, test, and refine a device that converts energy from one form to another. The PE uses the following three elements of the three-dimensional standards: (1) Constructing Explanations and Designing Solutions (i.e., SEP), (2) Conservation of Energy and Energy Transfer (i.e., DCI), and (3) Energy and Matter (i.e., CCC).

Part A requires students to demonstrate their knowledge of how energy is stored, transferred, or used within the system. In this item cluster, they must know how a battery, wires, and a speaker work within the circuit. This aligns with the DCI and the CCC.

Part B requires students to design and test designs that use electricity to produce a sound. This aligns with the DCI (how changes in current influence the production of sound) and the SEP (designing and testing solutions to a design problem).

Part C requires students to compare their designs with some criteria and constraints. This aligns with the SEP (designing and testing solutions) and the CCC (energy can be transferred in various ways and between objects). The answer for Part C is directly determined by how the student completes Part B. If all of the trials the student runs in Part B meet the given criteria, then all of those must be selected to be considered as correct in Part C. Therefore, there are multiple different ways to get this item correct.

Part D requires students to make a prediction from the evidence that they generated in Part B. This part is aligned to all three dimensions. The student has used their designs and information (representing SEP) from Part B to show how energy is transferred between objects (representing CCC) and specifically how increasing the current changes the volume (representing DCI).

Like Part C, Part E is dependent on Part B. Students are determining which trials support the prediction that they made in Part D. This part, combined with Part D and Part B, addresses all three dimensions of the PE.

The next big challenge for the MOU states was to properly score these item clusters so that all evidence of understanding the PEs and three dimensions could be collected. It was determined that scoring assertions would be the best way to capture and score student responses on item clusters. Scoring assertions were evidence statements that related specific features from the student response to skills and knowledge that were tested (of which they provide evidence). The use of these assertions in scoring created a direct linkage between what the student does and the inferences about the skills and knowledge that the student's response supports. This approach provided a physical embodiment of evidence-centered design, Mislevy and Haertel's well-regarded approach to cognitive measurement (Mislevy & Haertel, 2006). This also provided a structure for ensuring and reviewing alignment during test development and a clear explanation of what was measured, how it was measured, and why it was measured when tests were scored and reported.

By inspecting the student response for every meaningful piece of student input, more information about student skills and knowledge can be harvested than in a single interaction. In fact, evidence for some scoring assertions may derive from two or more interactions within an item cluster. This may happen if one interaction is dependent on another interaction, allowing for multiple solution paths. This is one of the primary reasons that scoring assertions within item clusters can show deeper cognitive understanding and higher-order thinking that is required of the three-dimensional science standards.

Each of the parts in an item cluster likely has one or more scoring assertions where student skills and knowledge are being collected. The scoring mechanism has the capability to focus on one interaction, one part, or to focus across multiple interactions and parts as determined by the item writers, subject-matter expert (SME) reviewers and performance expectations. All permutations and combinations of measurable actions can be captured with scoring assertions.

The example item cluster from this section has seven assertions. Each scoring assertion is described in detail in Figure 3.

Figure 3. Example of NGSS Scoring Assertions

Your response earned **7** points of a possible **7**

Score Rationale	
When asked to describe what is happening to the energy for the battery when the doorbell is turned on, the student selected "energy is stored" or "energy is transferred." This provides some evidence of an ability to complete a causal chain explaining how energy can be transferred via electric current to produce light, sound, heat, and /or motion.	✓
When asked to describe what is happening to the energy of the wires when the doorbell is turned on, the student selected "energy is transferred." This provides some evidence of an ability to complete a causal chain explaining how energy can be transferred via electric current to produce light, sound, heat, and /or motion.	✓
When asked to describe what is happening to the energy of the speaker when the doorbell is turned on, the student selected "electrical energy is converted to sound energy." This provides some evidence of an ability to complete a causal chain explaining how energy can be transferred via electric current to produce light, sound, heat, and /or motion.	✓
The student ran at least two trials and ran at least one trial in which they selected components of a doorbell that produced "Loud" or "Very Loud" sound and that included components that cost less than $40. This provides some evidence of an ability to select characteristics to be manipulated while gathering information to determine the loudest, cost-effective doorbell.	✓
When asked to select the trial that met the criteria for being heard upstairs and cost less than $40, the student selected all trials from their simulation that produced "Loud" or "Very Loud" sound and cost less than $40. This provides some evidence of an ability to use given information to design and test a device that converts energy from one form to another.	✓
When asked to predict what will happen to the sound of the doorbell if the battery power increases, the student selected "The loudness of the doorbell will increase because more energy is stored in the battery." This provides some evidence of an ability to use an explanation to predict how the sound of an object changes, given a change in the conversion of stored energy.	✓
When asked to select the trials that support the relationship between the loudness of the doorbell and the power of the battery, the student selected two trials from the simulation in which the loudness was higher for the trial with a battery with more power. This provides some evidence of an ability to use evidence to support an inference.	✓

Assertion texts like the one shown in Figure 3 are written for every assertion in every item. They describe the correct response and what evidence should be provided by the student's response.

In the example item cluster, Part A has three assertions. Each one "provides some evidence of an ability to complete a causal chain explaining how energy can be transferred via electric current to produce light, sound, heat, and/or motion." The student must know something about electrical energy (DCI) and how it is transferred or used (DCI and CCC) to correctly respond. One assertion corresponds to each row in the table (i.e., one for Battery, one for Wires, and one for Speaker).

Part B has two assertions. The first "provides some evidence of an ability to select characteristics to be manipulated while gathering information to determine the loudest, most cost-effective doorbell." The second assertion "provides some evidence of an ability to use given information to design and test a device that converts energy from one form to another." The student must use their knowledge of how electrical energy is used and transferred (DCI) and how to design and test a design of a device using electricity (SEP) to correctly interact with Part B.

Part C has one assertion, as the student's selections are not independent of each other. The assertion "provides some evidence of an ability to use given information to design and test a device that converts energy from one form to another." The student must be able use generated evidence to support a design decision (SEP) about the transfer of energy (CCC). This assertion is pulling responses from both Parts B and C. This is precisely how item clusters and assertions can assess multiple dimensions and higher levels of complexity, as students are running their own experiments and analyzing the outcomes, no matter what those outcomes are.

Part D has one assertion. The assertion "provides some evidence of an ability to use an explanation to predict how the sounds of an object changes, given a change in the conversion of stored energy." This shows how the student must use elements from all three dimensions to respond correctly to this assertion. The student uses data from their generated designs and makes a prediction using that data to support their knowledge of energy and energy transformations.

Part E also has one assertion. The assertion "provides some evidence of an ability to use evidence to support an inference." In this case, it is an inference about the relationship between the available battery power and the loudness of the bell. Again, this scoring assertion is pulling information from three different parts (Parts B, D, and E).

While each part of the item, each interaction within the item, or each assertion may not be three-dimensional, the item cluster as a whole represents all three dimensions. It also provides an organized flow of cognition from scaffolding (Part A), through the engineering process (Parts B and C), to a conclusion and evidentiary support of the conclusion (Parts D and E).

The assertion text explains how a student responded to a given task and what that task shows evidence of. This allows us to ensure that items allow each student an opportunity to show what they know and what their knowledge, skills, and abilities show about their understanding of science and engineering.

Once the item cluster, along with interactions and scoring assertions, was fully developed, CAI and the group of states were able to begin item and test development in earnest.

This item development process was managed by CAI's Item Tracking System (ITS), which is an auditable content-development tool that enforces rigorous workflow and captures each item change and comment. Reviewers, including internal CAI staff and external stakeholders in committee meetings, can review items in ITS as they will appear to the student, with all accessibility features and tools.

The process begins with the definition of item specifications and continues with

- selection and training of item writers;
- writing and internal review of items;

- review by state personnel and stakeholder committees;
- markup for translation and accessibility features;
- field-testing; and
- post–field-test reviews.

Each step of test development plays a role in ensuring that the items can support the claims on which they will be based. Table 1 shows how each step contributes to that goal and describes the steps in more detail.

Table 1. Summary of How Each Step of Development Supports the Validity of Claims

Developmental Steps	Support Alignment to the Performance Expectations	Reduces Construct-Irrelevant Variance Through Universal Design	Expands Access Through Linguistic and Other Supports
Item specifications	Specifies item interactions, content limits, and guidelines for meeting task demands and levels of cognitive engagement requirements and adjusting difficulty.	Avoids the use of any item interactions with accessibility constraints and provides language guidelines. Allows for multiple response modes to accommodate different styles.	
Selection and training of item writers	Ensures that item writers have the background to understand the PEs and item specifications. Teaches item writers how to select item interactions for measurement and accessibility.	Training in language accessibility, bias, and sensitivity helps item writers avoid unnecessary barriers.	
Writing and internal review of items	Checks content alignment and evaluates and improves overall quality.	Eliminates editorial issues and flags and removes bias and accessibility issues.	
Markup that prepares items for translation and accessibility features		Adds universal features, such as text-to-speech (TTS) for science, that reduce barriers.	Adds TTS, braille, American Sign Language (ASL), translations, and glossaries.
Review by state personnel and stakeholder committees	Checks content and cognitive complexity alignment; evaluates and improves overall quality.	Flags sensitivity issues.	
Field-testing	Provides statistical checks on quality and flags issues.	Flags items that appear to function differently for subsequent review to identify issues.	May reveal usability or implementation issues with markup.
Post–field-test reviews	Provides final, more focused checks on flagged	Provides final, focused review on items flagged for	

Developmental Steps	Support Alignment to the Performance Expectations	Reduces Construct-Irrelevant Variance Through Universal Design	Expands Access Through Linguistic and Other Supports
	items. Rubric validation ensures that scoring reflects PEs.	differential item functioning (DIF).	

2.2 Item Specifications

CAI is working with a group of states and territory, psychometricians, and science experts, including the authors of the Next Generation Science Standards (NGSS), to develop powerful innovative solutions to the challenges of measuring three-dimensional science standards based on the National Research Council's *A Framework for K–12 Science Education* published in 2012. Twelve states (Connecticut, Hawaii, Idaho, Montana, New Hampshire, North Dakota, Oregon, Rhode Island, South Dakota, Utah, West Virginia, and Wyoming) and one U.S. territory (the U.S. Virgin Islands) have participated in this initiative. This collaboration has yielded item specifications for PEs, sample item clusters for some specifications, and hundreds of science item clusters and stand-alone items. Under this collaboration, utilizing guidelines for item specifications proposed by WestEd in collaboration with the Council of Chief State School Officers (CCSSO), state and territory members, and content experts (CCSSO, 2015), states and one U.S. territory jointly developed item specifications.

Item specifications are documents designed to guide item writers as they craft test questions and stakeholders as they review those items. These specifications are intended to serve as a roadmap for writers to facilitate the creation of items that are properly aligned to the three dimensions comprising each science standard, and which together form coherent item clusters and stand-alone items. Science item specifications include the following elements:

- **Performance Expectation.** The PE provides the unique identifier and the text for the PE..
- **Dimensions.** This identifies the CCCs, SEPs, and DCIs that the PE assesses.
- **Clarifications and Content Limits.** This section delineates the specific content that the PE measures and the parameters in which items must be developed to assess the PE accurately, including the lower and upper complexity limits of items. Specifically, content limits refine the intent of the PE and provide limits of what may be asked of test takers. For example, content limits may identify the specific formulae that students are expected to know or not know.
- **Science Vocabulary.** This section identifies the relevant technical words that students are expected to know, and related words that they are explicitly not expected to know. These categories should not be considered exhaustive, as the boundaries of relevance are ambiguous and the list is limited by the imagination of the writers.

- **Content/Phenomena.** This section provides examples of the types of phenomena that would support the effective items related to the PE in question. In general, these are guideposts, and item writers seek comparable phenomena, rather than drawing on those within the documents.
- **Task Demands.** In this section, the PEs and associated evidence statements are broken down into specific task demands aligned to each PE. Task demands denote the specific ways in which students will provide evidence of their understanding of the concept or skill. Specifically, the task demands identify the types of interactions and activities that item writers should employ. Each item should be clearly linked to one or more of the task demands, and the verbs guide the types of interactions writers might employ to elicit the student response.

Table 2 provides a sample of the item specifications developed by content experts for a middle school Life Sciences PE.

Table 2. Sample Science Item Cluster Specifications for Middle School Life Sciences Performance Expectation

Performance Expectation	**MS-LS1-1[a]** Conduct an investigation to provide evidence that living things are made of cells; either one cell or many different numbers and types of cells.		
Dimensions	**Planning and Carrying Out Investigations** • Conduct an investigation to produce data to serve as the basis for evidence that meets the goals of an investigation.	**LS1.A: Structure and Function** • All living things are made up of cells, which is the smallest unit that can be said to be alive. An organism may consist of one single cell (unicellular) or many different numbers and types of cells (multicellular).	**Scale, Proportion, and Quantity** • Phenomena that can be observed at one scale may not be observable at another scale.
Clarifications and Content Limits	**Clarification Statements** • Emphasis is placed on developing evidence that living things are made of cells, distinguishing between living and non-living things, and understanding that living things may be made of one cell or many varying cells. **Content Limits** • Students do not need to know the following: ○ The structures or functions of specific organelles or different proteins ○ Systems of specialized cells ○ The mechanisms by which cells are alive ○ Specifics of DNA and proteins or of cell growth and division ○ Endosymbiotic theory ○ Histological procedures		
Science Vocabulary Students Are Expected to Know	Multicellular, unicellular, cell, tissue, organ, system, organism hierarchy, bacteria, colony, yeast, prokaryote, eukaryote, magnify, microscope, DNA, nucleus, cell wall, cell membrane, algae, chloroplast(s), chromosome, cork		

Science Vocabulary Students Are Not Expected to Know	Differentiation, mitosis, meiosis, genetics, cellular respiration, energy transfer, RNA, protozoa, amoeba, histology, protists, archaea, nucleoid, plasmid, diatoms, cyanobacteria
Phenomena	
Context/ Phenomena	Some example phenomena for MS-LS1-1 include the following: • Plant leaves and roots have tiny, box-like structures that can be seen under a microscope. • Small creatures can be seen swimming in samples of pond water viewed through a microscope. • Different parts of a frog's body (e.g., muscles, skin, tongue) are observed under a microscope, and are seen to be composed of cells. • One-celled organisms (e.g., bacteria, protists) perform the eight necessary functions of life, but nothing smaller has been seen to do this. • Swabs from the human cheek are observed under a microscope. Small cells can be seen.
This Performance Expectation and Associated Evidence Statements Support the Following Task Demands	
Task Demands	
1. Identify from a list the materials/tools, including distractors, needed for an investigation to find the smallest unit of life (cell).	
2. Identify the outcome data that should be collected in an investigation of the smallest unit of living things.	
3. Evaluate the sufficiency and limitations of data collected to explain that the smallest unit of living things is the cell.	
4. Make and/or record observations about whether the sample contains cells.[b]	
5. Interpret and/or communicate data from the investigation to determine if a specimen is alive.	
6. Construct a statement to describe the overall trend suggested by the observed data.	

[a]MS-LS1-1 is the PE code for Middle School Life Sciences 1-1.

[b]Denotes task demands deemed appropriate for use in stand-alone item development.

The specifications help test developers create item clusters and stand-alone items that will support a range of difficulty while remaining at grade level, furthering the goal of measuring the full range of performance found in the population.

The assertions provide evidence that the student is completing specific aspects of the PEs. Each assertion is reviewed during development and at all state and committee reviews to ensure that students are doing what is stated in the assertion. The assertion links the student response to an interaction to the performance expectation.

2.3 Selection and Training of Item Writers

All item writers developing science items at CAI have at least a bachelor's degree, and many bring teaching experience. All item writers are trained in

- the principles of universal design;
- the appropriate use of item interactions; and
- the science item specifications.

Key materials are shown in Appendix A, Item Writer Training Materials, Appendix B, Item Specifications Grade 3 through High School, and Appendix C, Style Guide for Science Items. These include

- CAI's language accessibility, bias, and sensitivity (LABS) guidelines;
- a training (presented using Microsoft PowerPoint) for the appropriate use of item interactions;
- item specification for science for grades 3 through high school; and
- style guide for science items.

2.4 INTERNAL REVIEW

CAI's test development structure uses highly effective units organized around each content area. Unit directors oversee team leaders who work with team members to ensure item quality and adherence to best practices. All team members, including item writers, are content-area experts. Teams include senior content specialists who review items before the client review and provide training and feedback for all content-area team members.

ICCR and MOU science items undergo a rigorous, multi-level internal review process before they are sent to external review. Staff members are trained to review items for both content and accessibility throughout the process. A sample item review checklist that CAI test development uses is included in Appendix 2-D, Item Review Checklist. The ICCR and MOU science internal review cycle includes the following phases:

- Preliminary Review
- Scoring Entry and Review
- Content Review One
- Edit Review
- Senior Review

2.4.1 Preliminary Review

Team leads or senior content staff conduct Preliminary Review. Sometimes, Preliminary Review is conducted in a group setting, led by a senior test developer. During the Preliminary Review process, team leads or senior content staff analyze items to ensure the following:

- The item aligns with the PE, including the listed SEP, DCI, and CCC. The item matches the item specification for the skills and knowledge being assessed. The item specification contains clarifying statements, content limits, and task demands, as well as knowledge, skills, and abilities that the PE is intended to assess.
- The item is based on a quality scientific phenomenon (i.e., it assesses something in a reasonable way, and it is a discrete observation that grounds a scenario, which allows

for the assessment of something worthwhile in a meaningful way). A quality phenomenon is one that is natural, observable (even with instrumentation), and focused on a specific event, not a general category of similar events (e.g., the effects of Hurricane Katrina, not hurricanes in general).

- The item aligns appropriately with the task demands. Task demands are statements about what a student is expected to do with a phenomenon.
- The vocabulary used in the item is appropriate for the grade and subject matter. Most non-technical language is two grade levels below the testing grade to ensure that language is not a construct-irrelevant issue.
- The item considers language accessibility, bias, and sensitivity.
- The content is accurate and straightforward.
- The graphic and stimulus materials are necessary to answer the question. The phenomenon is described in the stimulus. Graphics are necessary and contain only the relevant information.
- The item follows the approved style guide.
- The stimulus is clear, concise, and succinct (i.e., it contains enough information to convey what is being asked, it is stated positively, and it does not rely on negatives—such as *no, not, none,* or *never*—unless necessary).

For selected-response item interactions, test developers also check to ensure that the set of response options are

- as succinct and short as possible (without repeating text);
- parallel in structure, grammar, length, and content;
- sufficiently distinct from one another;
- all plausible (but with only one correct option); a plausible distractor is one that is related to the item, but contains a misconception, a logical error, or a pattern of thinking that a student might have, but is incorrect; and
- free of obvious or subtle cuing.

2.4.2 Scoring Entry and Review

During Scoring Entry, the item writer inputs the machine scoring for review by the team lead or senior staff before the Content Review One level. This step is separate from Preliminary Review to allow senior staff to suggest changes to the interaction at Preliminary Review without requiring the writer to overhaul the scoring they already created, ensuring that the scoring is entered once, streamlining the process. This step also allows senior staff to ensure that the scoring suggested by the writer at Preliminary Review is appropriate. At this level, scoring is analyzed to ensure the following criteria:

- The scoring works as intended (i.e., the student gets a point for ALL correct responses and no points for ALL incorrect responses).
- The student receives a point for every unique piece of information they reveal about their understanding through their responses.
- Dependent scoring between and within interactions is captured.
- The way in which the scoring is set up is unambiguous and matches the questions asked (i.e., if we ask students to round a number to a certain decimal place, we score accordingly).

Senior staffs approve the intent of the scoring from the Preliminary Review. At the Scoring-Entry level, item writers input the approved scoring. After that, senior staff check the functionality of the scoring. Once the scoring is determined to be working correctly, senior staff sign-off on the item and move it to Content Review One.

Senior staff are recruited based on experience and time in the assessment field. Senior staff are the reviewers of the intent of scoring because of their experience and knowledge of assessment, the expectations of the clients, and their understanding of student responses.

2.4.3 Content Review One

Content Review One is conducted by a senior content specialist who was not part of the Preliminary Review. This reviewer carefully examines each item based on all the criteria identified for Preliminary Review. He or she also ensures that the revisions made during the Preliminary Review did not introduce errors or content inaccuracies. This reviewer approaches the item by combining their expertise in test development while engaging from the perspective of potential clients.

2.4.4 Edit Review

During Edit Review, editors have the following four primary tasks:

1. Editors perform basic line editing for correct spelling, punctuation, grammar, and mathematical and scientific notation, ensuring consistency of style across the items.
2. Editors ensure that all items are accurate in content. Editors compare reading passages against the original publications to ensure that all information is internally consistent across stimulus materials and items, including names, facts, or cited lines of text that appear in the item. They ensure that the keys and all information in the item are correct. Keys are the correct answers to interactions. Information refers to the phenomena and the science content. For items with mathematical tasks, editors perform all calculations to ensure accuracy.
3. Editors review all material for fairness and language accessibility issues.
4. Editors confirm that items reflect the accepted guidelines for good item construction. In all items, they ensure that language is simple, direct, and free of ambiguity with minimal verbal difficulty. Editors confirm that a problem or task and its stem are clearly defined and concisely worded with no unnecessary information. For multiple-

choice interactions, editors check that options are parallel in structure and fit logically and grammatically with the stem and that the key answer (i.e., correct answer) accurately and correctly answers the question as posed, that the answer is not inappropriately obvious, and that it is the only correct answer to an item among the distractors. For constructed-response interactions, editors review the rubrics for appropriate style and grammar.

2.4.5 Senior Review

By the time a science item arrives at Senior Review, both content reviewers and editors have thoroughly vetted it. Senior reviewers (in particular, senior content specialists) look at the item's entire review history, ensuring that all the issues identified in that item have been adequately addressed. Senior reviewers verify the overall content of each item, confirming its accuracy, alignment to the PE, and consistency with expectations for the highest quality. They check whether the scoring is working as intended and scoring assertions adequately address the evidence the student provides with each type of response.

Some examples of questions from the internal Review Checklist are listed below. These are the questions that reviewers ask of the item to ensure that it is three dimensional and properly aligned to the PE. A similar checklist is used at earlier stages.

- Is the phenomenon based on a specific real-world scenario and focused enough to get the student to investigate what the PE intends for them to investigate (i.e., the students' application of the Practice in the context of the DCI and CCC as intended by the PE is sufficient to make sense of the phenomena)?
- What information should the student already have before starting the cluster (DCI knowledge)?
- Cluster Task Statement: Does it align to the focus and intent of the PE?
- Does the interaction require the student to demonstrate the science practice and/or content that the PE is assessing them on?
- Do the interactions align with the task demands?

2.5 Review by State Personnel and Stakeholder Committees

All science items undergo an exhaustive external review process. Items in the Shared Science Assessment Item Bank were reviewed by content experts in one or several states or territory and reviewed and approved by multiple stakeholder committees that evaluated them for both content and bias/sensitivity.

2.5.1 State Review

After items have been developed for a state or territory participating in the MOU, content experts from the state or territory that owns the item review any eligible items before committee review. At this stage in the review process, clients can request edits, such as wording edits, scoring edits, alignment changes, or task demand updates. A CAI science content expert reviews all client-requested edits concerning the science item specifications, other clients' requests, and existing items in the bank to determine whether the requested edits will be made. At this stage, clients have the option to present these items to the committee (based on the edits made) or withhold them from committee review.

ICCR items are reviewed by at least three individuals from one or more states in the MOU. The states or territory provide feedback on the ICCR items, and CAI science leadership gathers suggestions and makes edits that improve the ICCR item. Not all suggestions are implemented, as CAI owns these items. Further, most MOU states accept or reject ICCR and MOU items (as they appear at the time) to be presented to their committees. Some clients skip this step and allow CAI to review all items with their committees before reviewing them. These items can be either set for field-testing in a future administration or become a part of the locked operational pool.

2.5.2 Content Advisory Committee Reviews

During the Content Advisory Committee (CAC) reviews, items are reviewed for content accuracy, grade-level appropriateness, and alignment to the PE. CAC members are typically grade-level and subject-matter experts. During this review, educators also ensure that the scoring assertions clearly identify what is being scored as correct and give credit where they should (refer to Section 2.7.1, Rubric Validation). Before the CAC review begins, CAI provides a presentation on the three-dimensional science standards, the item development process, the CAI systems that will be used in the review, and how to review the items for content. Appendix 2-E, Content Advisory Committee Review Training Slides contains of the slides used during the CAC review training.

Items developed for each state under the MOU are reviewed by the state that owns those items. ICCR items are reviewed by the CAC of one or more states. In most cases, items are reviewed by multiple state committees before their field-testing or operational use.

In 2023, all MOU states engaged in a single CAC event in which participants from multiple states reviewed items. The items were edited and returned to their respective owner states for final approval.

A summary of the 2022-2023 committee meetings is presented in Table 3, with additional details about the participants in Appendix 2-F, Content Advisory Committee Participant Details. This appendix also contains detailed information about the participants of CAC meetings of previous years.

Table 3. Summary of the 2022-2023 Content Advisory Committee Meetings

State/ Item Bank	Meeting	Number of Committee Members	Number of Items Reviewed
Connecticut	June/ August 2022[a]	26	65
	July 2022	21	62
Hawaii	June/ August 2022[a]	25	46
	July 2022	9	45
	July 2022	9	306
ICCR	June/ August 2022[a]	12	121

State/ Item Bank	Meeting	Number of Committee Members	Number of Items Reviewed
Idaho	June/ August 2022[a]	14	12
	July 2022	5	244
Montana	June/ August 2022[a]	9	13
Oregon	July 2022	14	66
Rhode Island	October 2022	20	115
Utah	September 2022	28	111
West Virginia	June/ August 2022[a]	8	13
Wyoming	June/ August 2022[a]	9	37

[a]Items were reviewed in a combined Content Advisory Committee Meeting that included ICCR and MOU state-owned items. Items reviewed in the combined meetings are displayed by their respective state or bank of ownership.

2.5.3 Language Accessibility, Bias, and Sensitivity Committee Reviews

During bias and sensitivity reviews, stakeholders review items to check for issues that might unfairly impact students based on students' backgrounds. For example, some states include representatives from student populations such as Special Education, low vision, and the hearing impaired. Furthermore, diverse members of this committee represent the interests of students from various ethnic and economic backgrounds to ensure that test items are free of bias and sensitivity concerns. States try to ensure that all demographics are represented when providing committee members. For example, if a state has a Native American population, they will try to ensure that the Bias and Sensitivity Committee has Native American representation on the committee. Before the bias and sensitivity review begins, CAI provides a presentation on the three-dimensional science standards, the item development process, the CAI systems that will be used in the review, and how to review items for fairness. Appendix 2-E, Content Advisory Committee Review Training Slides contains the slides used during the bias and sensitivity review training.

During 2020–2022, due to the COVID-19 pandemic, CAI content experts reviewed 65 items that contained references to *virus*, *vaccine*, *bacteria*, *disease*, *infection*, and related words and phrases. One item of these 65 was rejected for sensitivity concerns.

In 2023, the MOU states were all involved in a single review process whereby participants from multiple states would review items. The items were edited and then returned to the owning state for final approval.

A summary of the committee meetings is presented in Table 4, with additional details about the participants in Appendix 2-H, Fairness Committee Participant Details. This appendix also contains detailed information about the participants of Fairness Committee meetings of previous years.

Table 4. Summary of the 2022-2023 Fairness Committee Meetings

State/ Item Bank	Meeting	Number of Committee Members	Number of Items Reviewed	Number of Items Rejected
Connecticut	June/ August 2022[a]	3	65	2
	August 2022	19	154	27
Hawaii	June/ August 2022[a]	6	46	0
	July 2022	9	45	0
	July 2022	9	306	0
ICCR	June/ August 2022[a]	7	121	3
Idaho	June/ August 2022[a]	4	12	0
Montana	June/ August 2022[a]	4	13	0
Oregon	July 2022	9	43	2
Rhode Island	October 2022	20	115	22
Utah	September 2022	28	111	12
West Virginia	June/ August 2022[a]	3	13	0
Wyoming	June/ August 2022[a]	6	37	0

[a]Items were reviewed in a combined Fairness Committee Meeting that included ICCR and MOU state-owned items. Items reviewed in the combined meetings are displayed by their respective state or bank of ownership.

2.5.4 Markup for Translation and Accessibility Features

After all approved state- and committee-recommended edits have been applied, the items are considered "locked" and ready for a portion of the accessibility tagging. Text-to-speech (TTS) tagging is applied prior to field-testing, while Spanish translations and braille are applied post-field-testing. Accessibility markup is embedded into each item as part of the item development process rather than as a *post-hoc* process applied to completed tests.

Accessibility markup, whether for translations or TTS, follows similar processes. One trained expert enters the markup, and then a second expert reviews the work and recommends changes, if necessary. If there is disagreement, a third expert is engaged to resolve the issue.

Currently, science items are tagged with TTS. Spanish translations, including Spanish TTS and braille, are available for a subset of items.

2.6 Field-Testing

A large pool of science field-test items was administered in the following nine states in spring 2018: Connecticut, Hawaii, New Hampshire, Oregon, Rhode Island, Utah, Vermont, West Virginia, and Wyoming. For three other states—Hawaii, Oregon, and Wyoming—items were embedded as field-test items in the legacy science test. Connecticut and Rhode Island conducted an independent field test in which all students participated but for which no scores were reported. In New Hampshire, Utah, Vermont, and West Virginia, an operational field test was administered.

In 2019, a second wave of field-test items was administered in the following nine states: Connecticut, Hawaii, Idaho, New Hampshire, Oregon, Rhode Island, Vermont, West Virginia, and Wyoming. In Hawaii, Idaho (elementary school), and Wyoming, unscored field-test items were added as a separate segment to the operational (scored) legacy science test. An independent field test in which students were administered a full set of items was conducted for a sample of Idaho middle schools. In Connecticut, New Hampshire, Oregon, Rhode Island, Vermont, and West Virginia, field-test items were administered as unscored items embedded within the operational items.

In 2021, a third wave of field-test items was administered in 12 states: Connecticut, Hawaii, Idaho, Montana, New Hampshire, North Dakota, Rhode Island, South Dakota, Utah, Vermont, West Virginia, and Wyoming. An independent field test, in which students were administered a full set of items, was conducted for Idaho and Montana. Unscored field-test items were added as a separate segment to the operational (scored) legacy science test for Wyoming. In the remaining nine states, field-test items were administered as unscored items embedded within the operational items.

In 2022, a fourth wave of field-test items was administered in 13 states and one U.S. territory: Connecticut, Hawaii, Idaho, Montana, New Hampshire, North Dakota, Oregon, Rhode Island, South Dakota, Utah, Vermont, West Virginia, Wyoming, and U.S. Virgin Islands. In all 13 states and the territory, field-test items were administered as unscored items embedded within the operational items.

In 2023, items were field-tested in 12 states and one U.S. territory: Connecticut, Hawaii, Idaho, Montana, New Hampshire, North Dakota, Oregon, Rhode Island, South Dakota, Utah, West Virginia, Wyoming, and U.S. Virgin Islands. Field-test items were administered as unscored items embedded within the operational items. CAI's field-test process is described in detail in Section 3.2, Field-Testing, in Volume 1 of this technical report.

2.7 Post–Field-Test Review

Following field-testing, items are subjected to a substantial validation process. This includes rubric validation and data review. These processes are described in Section 2.7.1, Rubric Validation, and Section 2.7.2, Data Review.

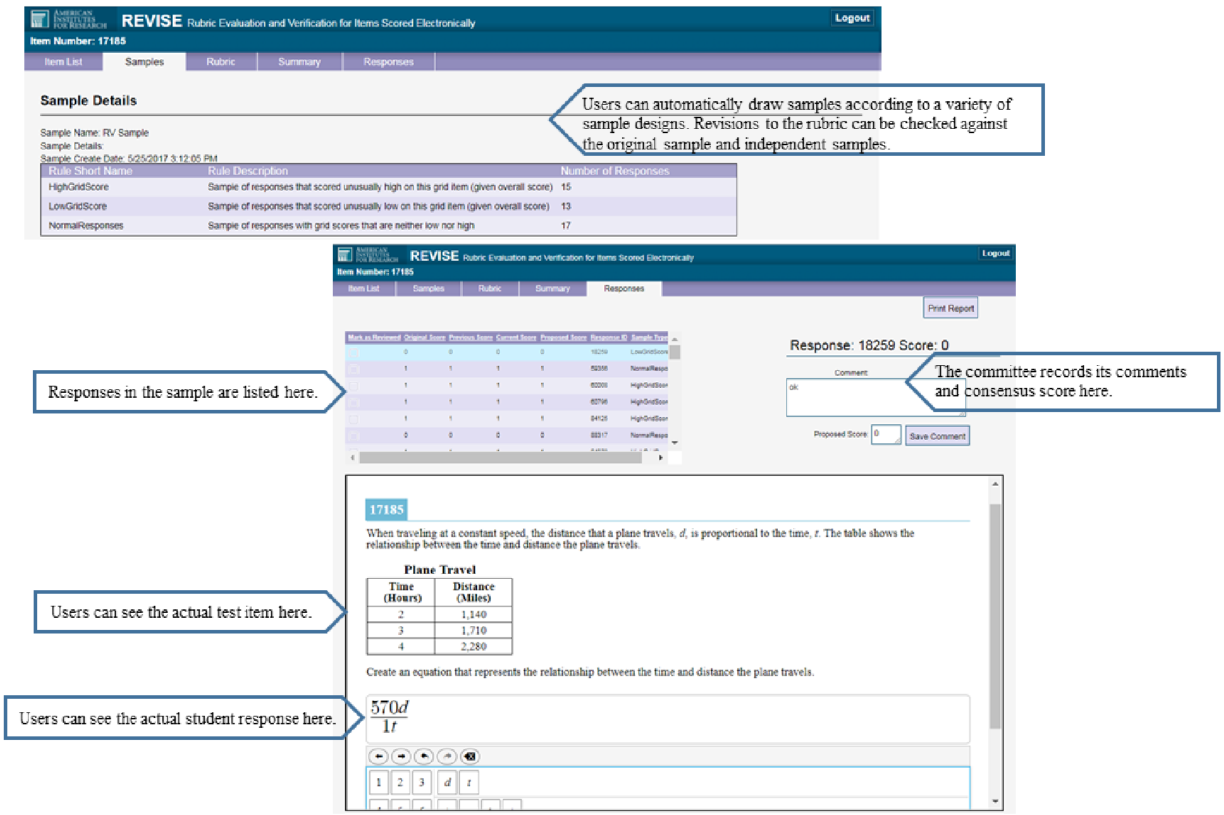
2.7.1 Rubric Validation

The validation process for the field-test items begins with rubric validation to verify and make any necessary revisions to the scoring rubrics. The rubric validation process occurs in two phases. During the first phase, CAI content experts work with the analysis team to prepare for the rubric validation meetings. The CAI content experts use the Rubric Evaluation and Verification for Items Scored Electronically (REVISE) system to generate student responses that are scientifically sampled to overrepresent responses most likely to have been mis-scored. Specifically, the sample overrepresents low-scored responses from otherwise high-scoring students and high-scored responses from otherwise low-scoring students. This process allows CAI to identify any potential scoring concerns before the rubric validation meeting, such as unanticipated (but accurate) responses, equivalent responses that were not originally considered, and responses that received credit but should not have (based on the content and the item rubric). At this point, the rubrics may be adjusted, and responses rescored.

The second phase of rubric validation involves committees of educators in each state. The committees review the response samples generated by CAI to make recommendations to either change or confirm the rubrics of each item. The committee recommendations are then discussed with the state of ownership to resolve any inconsistencies. The rubrics are then edited or confirmed based on this resolution.

Figure 4 illustrates the features provided by the REVISE system.

Figure 4. Features of the REVISE Software



After the rubric validation meetings, CAI staff apply the approved revisions to the rubrics, and any items rejected as part of the process are rejected in ITS. During this process, ITS archives critical information regarding the scoring certification completed during the rubric validation process. This includes recording any rubric changes made during the scoring decision meetings and the sign-off completed by the senior content expert once the rubric has been changed, rescoring the entire sample, and verifying that the final rubric functioned as intended.

Following rubric validation, all items are subject to statistical checks, and flagged items are presented in data review committees.

2.7.2 Data Review

Following rubric validation, all items are rescored and classical item statistics are computed for the scoring assertions, including item difficulty and item discrimination statistics, testing time, and differential item functioning (DIF) statistics. The states established standards for the statistics, and any items violating these standards are flagged for a second educator review. Even though the scoring assertions are the basic units of analysis to compute classical item statistics, the business rules to flag items for additional educator review are established at the item level because assertions cannot be reviewed in isolation. A common set of business rules is defined for all the states participating in the field test. The classical item statistics are computed on the data of the students testing in the state that owned the item. For ICCR items administered in spring 2023, the data from students testing in Connecticut, Idaho, New Hampshire, North Dakota, Oregon, Rhode Island, South Dakota, Utah, and West Virginia were combined (states that administered ICCR items and used either an independent field test or operational test).

Section 4 of Volume 1, Field-Test Classical Analysis, describes the statistical flags that designate items for data review. The flags are designed to highlight potential content weaknesses, miskeys, or possible bias issues. Committee members are taught to interpret these flags and are given guidelines for examining the items for content or fairness issues.

For each of the states participating in the MOU, flagged items owned by the state are reviewed by a data review committee. The composition of the data review committees generally includes content experts from the state's department of education or state educators (in this case, the state educators were science teachers) and are supported by CAI content experts. ICCR field-test items are taken to committee members from several states participating in the MOU. The outcomes are decided by CAI science content leadership, taking the committees' recommendations into consideration.

At the start of each state-owned item data review meeting, CAI staff leads participants in a training session to familiarize them with the item development process, the purpose of the data review committee and the data review process, and the meaning of the various flags. Committee members are taught to interpret the various flags and are given guidelines for examining the items for content or fairness issues. The training includes a group review of item cards, which detail specific item attributes (e.g., grade level and alignment to the science PEs, the content and rubric of the item, and various item statistics). A sample of the training materials used for these data review meetings is presented in Appendix 2-I, Sample Data Review Training Materials. Participants use an online environment via laptop computers to review the items and interact with them in a manner similar to that of students, and to view the statistics associated with each item.

The items are then reviewed by the participants who are most familiar with the particular grade (or grade-band) level and the items' content domain. CAI content specialists, who are also well versed in item statistics, facilitate the discussion in each room with CAI psychometricians available to answer questions as they arise. At the end of each meeting day, CAI content specialists meet with the state content specialists to review the committee recommendations and decide whether to accept or reject the item for inclusion in the operational pool. Items that were rejected become eligible for potential changes and additional field-test items.

Table 5 summarizes the data review committee meetings. Details, including the composition of each committee, are presented in Appendix 2-J, Data Review Committee Participant Details.

Table 5. Summary of Data Review Committee Meetings

State/ Item Bank	Meeting	Number of Committee Members	Item Type	Number of Items Reviewed	Number of Items Rejected
Connecticut	August 2018	29	Cluster	7	5
			Stand-Alone	11	6
	August 2019	29	Cluster	14	6
			Stand-Alone	39	11
	August 2021	19	Cluster	8	2
			Stand-Alone	43	10
	August 2022	15	Cluster	5	4
			Stand-Alone	14	2
	September 2023	12	Cluster	11	4
			Stand-Alone	32	17
Hawaii	August 2018	18	Cluster	7	1
			Stand-Alone	25	2
	August 2019	18	Cluster	17	5
			Stand-Alone	20	8
	August 2021[a]	25	Cluster	6	0
			Stand-Alone	20	8
	August 2022[a]	12	Cluster	11	2
			Stand-Alone	38	6
	August 2023[a]	15	Cluster	3	2
			Stand-Alone	23	3
ICCR	July 2018	18	Cluster	33	2
			Stand-Alone	51	6
	August 2019[b]	–	Cluster	0	1
			Stand-Alone	43	2
	August 2021[a]	25	Cluster	11	2
			Stand-Alone	64	4
	August 2022[a]	20	Cluster	12	1
			Stand-Alone	56	13
	August 2023[a]	19	Cluster	12	1
			Stand-Alone	42	8
Idaho	August 2019	10	Cluster	4	3
			Stand-Alone	8	3
	August 2021[a]	25	Cluster	26	1
			Stand-Alone	34	4
	August 2022[a]	8	Cluster	3	0
			Stand-Alone	1	0
	August 2023[a]	17	Cluster	2	0
			Stand-Alone	3	0
Montana	September 2021	4	Cluster	3	2
			Stand-Alone	14	2
	September 2022	5	Cluster	5	2
			Stand-Alone	12	1
	August 2023[a]	11	Cluster	2	1
			Stand-Alone	10	2
Multi-State Science	August 2018	–	Cluster	2	0
			Stand-Alone	7	6

State/ Item Bank	Meeting	Number of Committee Members	Item Type	Number of Items Reviewed	Number of Items Rejected
Assessment (Rhode Island and Vermont)	August 2019	–	Cluster	2	1
			Stand-Alone	12	3
	August 2021	–	Cluster	4	4
			Stand-Alone	14	5
	September 2022	–	Cluster	1	1
			Stand-Alone	10	6
Oregon	September 2018	11	Cluster	28	5
			Stand-Alone	16	1
	August 2019	4	Cluster	1	1
			Stand-Alone	7	6
	August 2022[a]	8	Cluster	11	2
			Stand-Alone	20	6
			Legacy Stand-Alone	*9*	*4*
	August 2023[a]	16	Cluster	9	1
			Stand-Alone	3	1
			Legacy Stand-Alone	*24*	*11*
Rhode Island	September 2023	–	Cluster	7	2
			Stand-Alone	10	4
South Dakota	September 2021	–	*Legacy Stand-Alone*	*15*	*0*
	September 2022	–	*Legacy Stand-Alone*	*4*	*1*
	September 2023	–	*Legacy Stand-Alone*	*6*	*2*
Utah	August 2018	16	Cluster	40	6
	September 2021	6	Cluster	11	3
	September 2022	13	Cluster	11	6
	September 2023	20	Cluster	6	0
West Virginia	July 2018	4	Cluster	3	1
			Stand-Alone	0	0
	September 2019	4	Cluster	1	1
			Stand-Alone	6	5
	August 2021[a]	25	Cluster	1	1
			Stand-Alone	6	2
	August 2022[a]	9	Cluster	4	2
			Stand-Alone	6	2
	August 2023[a]	11	Cluster	2	1
			Stand-Alone	10	2
Wyoming	October 2018	12	Cluster	6	1
			Stand-Alone	10	5
	August 2019	10	Cluster	4	3
			Stand-Alone	12	2
	August 2021[a]	25	Cluster	3	1
			Stand-Alone	13	3
	August 2022[a]	12	Cluster	2	0
			Stand-Alone	17	3
	August 2023[a]	17	Cluster	3	0
			Stand-Alone	5	1

Note. MSSA, Rhode Island, and South Dakota-owned items were reviewed by Rhode Island Department of Education and Vermont Agency of Education science content experts, the Rhode Island Department of Education, and the South Dakota Department of Education, respectively.

[a]Combined Item Data Review Meetings were conducted for multiple states in 2021, 2022, and 2023 (184 items were reviewed in the combined meeting format for Hawaii, Idaho, West Virginia, Wyoming, and ICCR items in 2021; 181 items were reviewed in the combined meeting format for Hawaii, Idaho, Oregon, West Virginia, Wyoming, and ICCR items in 2022, and 129 items were reviewed in the combined meeting format for Hawaii, Idaho, Montana, Oregon, West Virginia, Wyoming, and ICCR items in 2023). In 2021, 25 committee members took part in the combined Item Data Review Meetings; in 2022, 38 committee members participated in the combined Item Data Review Meetings, and in 2023, 41 committee members participated in the combined Item Data Review Meetings. Items reviewed in the combined meetings are displayed by their respective state or bank of ownership.
*b*During the summer 2019, ICCR field-test items underwent committee review in Connecticut, Hawaii, and Idaho.

3. Shared Science Assessment Item Bank Summary

Tests based on *A Framework for K–12 Science Education* (National Research Council, 2012) adopt a three-dimensional conceptualization of science understanding, including crosscutting concepts (CCCs), science and engineering practices (SEPs), and disciplinary core ideas (DCIs). Accordingly, the new science assessments are composed mostly of item clusters representing a series of interrelated student interactions directed towards describing, explaining, and predicting scientific phenomena. Some stand-alone items are added to increase the coverage of the test without increasing the testing time or testing burden.

CAI has built the Shared Science Assessment Item Bank in partnership with multiple states and one U.S. territory. The science item bank is robust and has been constructed to support multiple statewide science assessments. As described earlier, science items are written to the three-dimensional science standards. The Shared Science Assessment Item Bank comprises ICCR items and items developed for specific states, which are all shared with MOU partner states. These items follow the same specifications, test development processes, and review processes.

In 2018, CAI field-tested more than 540 item clusters and stand-alone items, of which 451 (including items from all sources) were accepted and made available as operational items in 2019. In 2019, 347 item clusters and stand-alone items were field-tested, of which 268 were accepted and made available as operational items in 2020. In 2021, CAI field-tested 545 item clusters and stand-alone items, of which 458 have passed rubric validation and item data review. In 2022, CAI field-tested 471 item clusters and stand-alone items, of which 403 have passed rubric validation and item data review. In 2023, CAI field-tested 348 item clusters and stand-alone items, of which 288 have passed rubric validation and item data review.

Each state or territory using the Shared Science Assessment Item Bank selects items that are appropriately aligned and have passed required reviews (as described in Section 2, Item Development Process That Supports Validity of Claims) for use on its statewide assessment. The Shared Science Assessment Item Bank continues to grow as participating states and territory continue to field-test new items. Participating states and territory collectively share the items and agree to field-test new items each year.

3.1 CURRENT COMPOSITION OF THE SHARED SCIENCE ASSESSMENT ITEM BANK

The Shared Science Assessment Item Bank contains item clusters and stand-alone items. Item clusters represent a series of interrelated student interactions directed toward describing, explaining, and predicting scientific phenomena. Item clusters can comprise several item parts requiring the student to interact with the item in various ways. In addition, shorter items (stand-alone items) are included to increase the coverage of the assessments without also increasing testing time or testing burden.

Within each item (including both item clusters and stand-alone items), a series of explicit assertions is made about the knowledge and skills that a student has demonstrated based on specific features of the student's responses across multiple interactions. For example, a student may correctly graph data points indicating that they can construct a graph showing the relationship between two variables, but they may make an incorrect inference about the relationship between the two variables, therefore not supporting the assertion that the student can interpret relationships expressed graphically.

Each year, before development begins, CAI reviews the bank for the state and for the MOU as a whole against the state operational blueprint. The number of eligible items (stand-alone and cluster items) are compared to the requirements of the blueprint. The target numbers are five times the maximum number of clusters and 10 times the maximum number of stand-alone items per Performance Expectation (PE) in the blueprint. If the blueprint has a maximum of one cluster per PE and one stand-alone item per PE then the total target is 15 (i.e., five clusters and 10 stand-alone items). This comparison provides information on areas that need further development. In addition to this target analysis, there are other conditions that might warrant additional item development. These things may include difficulty of the items, releasing operational items to a practice or interim test, a changed blueprint, or overexposed items. CAI creates an Item Development Plan (IDP) to address these areas, as possible, within the limits of the development for that cycle. The IDP is reviewed and approved by the state partner.

Table 6 lists and describes the science interaction types. Examples of various interaction types can be found in Appendix 2-K, Sample Item Interactions.

Table 6. Science Interaction Types and Descriptions

Interaction Type	Associated Sub-Types	Description
Choice	Multiple-Choice	Traditional multiple-choice interaction allows students to select a single option from a list of possible answer options.
	Multi-Select	Traditional multi-select interaction (checkboxes) allows students to select one or more options from a list of possible answer choices.
Text Entry	Simple Text Entry	Students type a response in a text box.
	Embedded Text Entry	Students type their response in one or more text boxes that are embedded in a section of read-only text.

Interaction Type	Associated Sub-Types	Description
	Natural Language	Students are directed to provide a short, written response.
	Extended-Response	Students are directed to provide a longer, written response in the form of an essay.
Table	Table Match	Interaction allows students to check a box to indicate if the information from a column header matches information from a row header.
	Table Input	Interaction solicits students to complete tabular data.
Edit Task	Edit Task	Students click a word and replace it with another word that they type to revise a sentence.
	Edit Task with Choice	Students click a word or phrase and select the replacement from several options.
	Edit Task Inline Choice	Drop-down menus are placed throughout the text, and students select an option to complete the text.
Hot-Text	Selectable	Selectable hot-text interactions require students to select one or more text elements in the response area.
	Re-orderable	Re-orderable hot-text interactions require students to click and drag hot-text elements into a different order.
	Drag-from-Palette	Drag-from-Palette hot-text interactions require students to drag elements from a palette into the available blank table cells or "gaps" (text boxes) in the response area.
	Custom	Custom hot-text interactions combine the functionality of the other hot-text interaction sub-types. Students responding to a custom hot-text interaction may need to select text elements, rearrange text elements, and/or drag text elements from a palette to blank table cells or drop targets in the response area.
Equation	N/A	Equation interactions require students to enter a response into input boxes. These boxes may stand alone, or they may be in line with text or embedded in a table. The equation interaction may have an on-screen keypad which may consist of special mathematic characters. Students may also enter their response via a physical keyboard.
Grid	Grid	Grid interactions require students to enter a response by interacting with a grid area in the answer space. The student may be required to draw a line or shape, plot a point, or create a graph. The student may also drag and drop or click on selectable hot-spots.
	Hot-Spot	Hot-spot interaction sub-types facilitate grid interactions with specific hot-spot functionality. These interactions require students to select hot-spot regions in the grid area.
	Graphic Gap Match	Graphic gap match interactions facilitate grid interactions with specific drag-and-drop functionality. These interactions require students to drag image objects from a palette to specified regions (gaps) in the grid area.
Simulation	N/A	Simulation interactions allow students to investigate a phenomenon by selecting variables to get output data. Some simulations are accompanied by animations.

Table 7 - Table 11 present the number of items in the Shared Science Assessment Item Bank available for use in the spring 2023 statewide assessments. Appendix 2-L, Shared Science Assessment Item Bank, provides the items available within the bank by grade band, PE, and origin.

Table 7. Spring 2023 Shared Science Assessment Operational and Field-Test Item Bank

Grade Band and Item Type	ICCR Items	Montana Items	MOU Items[a]	Total Bank Items
Elementary School	**174**	**42**	**461**	**677**
Cluster	66	13	268	347
Stand-Alone	108	29	193	330
Middle School	**188**	**33**	**471**	**692**
Cluster	74	13	249	336
Stand-Alone	114	20	222	356
Total	**362**	**75**	**932**	**1369**

[a]Other MOU states include Connecticut, Hawaii, Idaho, Oregon, Rhode Island, Utah, West Virginia, and Wyoming.

Table 8. Spring 2023 Shared Science Assessment Operational Item Bank

Grade Band and Item Type	ICCR Operational Items	Montana Operational Items	MOU Operational Items[a]	Total Bank Operational Items
Elementary School	**139**	**30**	**382**	**551**
Cluster	49	9	229	287
Stand-Alone	90	21	153	264
Middle School	**147**	**24**	**385**	**556**
Cluster	53	8	216	277
Stand-Alone	94	16	169	279
Total	**286**	**54**	**767**	**1107**

[a]Other MOU states include Connecticut, Hawaii, Idaho, Oregon, Rhode Island, Utah, West Virginia, and Wyoming.

Table 9. Spring 2023 Shared Science Assessment Field-Test Item Bank

Grade Band and Item Type	ICCR Field-Test Items	Montana Field-Test Items	MOU Field-Test Items[a]	Total Bank Field-Test Items
Elementary School	**35**	**12**	**79**	**126**
Cluster	17	4	39	60
Stand-Alone	18	8	40	66
Middle School	**41**	**9**	**86**	**136**
Cluster	21	5	33	59
Stand-Alone	20	4	53	77

Grade Band and Item Type	ICCR Field-Test Items	Montana Field-Test Items	MOU Field-Test Items[a]	Total Bank Field-Test Items
Total	**76**	**21**	**165**	**262**

[a]Other MOU states include Connecticut, Hawaii, Idaho, Oregon, Rhode Island, Utah, West Virginia, and Wyoming.

Table 10. Spring 2023 Shared Science Assessment Operational and Field-Test Item Bank by Science Discipline

Grade Band	Science Discipline	Item Type	ICCR Items	Montana Items	MOU Items[a]	Total Bank Items[b]
Elementary School	Earth and Space Sciences	Cluster	23	5	86	114
		Stand-Alone	31	8	60	99
	Life Sciences	Cluster	21	4	77	102
		Stand-Alone	37	10	55	102
	Physical Sciences	Cluster	22	4	105	131
		Stand-Alone	40	11	78	129
Middle School	Earth and Space Sciences	Cluster	22	4	62	88
		Stand-Alone	30	9	65	104
	Life Sciences	Cluster	29	5	98	132
		Stand-Alone	50	5	82	137
	Physical Sciences	Cluster	23	4	82	109
		Stand-Alone	34	6	74	114
Total			**362**	**75**	**924**	**1361**

[a]Other MOU states include Connecticut, Hawaii, Idaho, Oregon, Rhode Island, Utah, West Virginia, and Wyoming.
[b]Count excludes eight MOU items that do not align to the NGSS.

Table 11. Spring 2023 Shared Science Assessment Operational and Field-Test Item Bank by Disciplinary Core Idea

Grade Band	Science Discipline	Disciplinary Core Idea	ICCR Items	Montana Items	MOU Items[a]	Total Bank Items[b]
Elementary School	Earth and Space Sciences	ESS1	11	5	39	55
		ESS2	19	4	70	93
		ESS3	24	4	37	65
	Life Sciences	LS1	22	5	49	76
		LS2	6	2	28	36
		LS3	7	2	20	29
		LS4	23	5	35	63
	Physical Sciences	PS1	14	2	45	61
		PS2	22	7	36	65
		PS3	19	3	65	87
		PS4	7	3	37	47
Middle School	Earth and Space Sciences	ESS1	14	4	36	54
		ESS2	22	4	43	69
		ESS3	16	5	48	69
	Life Sciences	LS1	27	2	65	94
		LS2	26	5	45	76
		LS3	6	2	17	25
		LS4	20	1	53	74
	Physical Sciences	PS1	18	5	46	69
		PS2	8	3	48	59
		PS3	20	1	37	58
		PS4	11	1	25	37
Total			**362**	**75**	**924**	**1361**

[a]Other MOU states include Connecticut, Hawaii, Idaho, Oregon, Rhode Island, Utah, West Virginia, and Wyoming. [b]Count excludes eight MOU items that do not align to the NGSS.

3.2 STRATEGY FOR ITEM BANK EVALUATION AND REPLENISHMENT

Both CAI and the participating MOU states continue to develop items to replenish and grow the science item bank. The general strategy for targeting item development involves gathering information from the following three sources:

Characteristics of released items to be replaced

1. Characteristics of items that are overused
2. Tabulations of content coverage and ranges of difficulty (to identify gaps in the bank)

Before a test goes live, simulations are used to fine-tune the parameters of the algorithm that govern the item selection in an adaptive test design. Among the many reports from the simulator are items that are seen by more than 20% of students. The characteristics of these items are the primary targets for development. Overused items become candidates for release in two years, once replacements have been introduced into the operational bank.

4. CONSTRUCTION OF THE MONTANA SCIENCE ASSESSMENT

4.1 TEST DESIGN

The Montana Science Assessment (MSA) was administered online in spring 2023 to students in grades 5 and 8 using an adaptive design. Appendix 2-M, Montana Science Assessment Item Pool, provides the 2023 item pool by grade band, performance expectation (PE), and origin. In an adaptive test, operational items are selected on the fly based on the performance of a student on past items while ensuring the test blueprint is followed for each individual student. An advantage of adaptive testing is that it can provide more precise scores for students with lower and higher proficiencies, in contrast to fixed forms and linear-on-the-fly tests (LOFTs) that are typically targeted to provide the best precision for students with medium proficiencies. Also, as opposed to a fixed form and a LOFT, every student has the potential to see a different set of items that adapt to the student's ability, thus offering a better testing experience.

Items are selected by an item-selection algorithm based on the content and information value. At any given point during the test, the content value of an item is determined by its contribution to meeting the blueprint, given the content characteristics of the items that have already been administered. During the test, the content value increases for items that exhibit features that have not met their designated minimum as the end of the test approaches. Conversely, the content value decreases for items with content features that met the minimum. The information value of an item is based on the item information function evaluated at the estimated proficiency. The proficiency estimate is updated throughout the test.

The adaptive item-selection algorithm is the same algorithm CAI uses to deliver ELA and mathematics tests, but with some modifications to make it suitable for using item clusters. Specifically, the proficiencies that are estimated during the test are computed under an item response theory (IRT) model that incorporates cluster effects. In order to avoid over-selection of items with many scoring assertions, the information of an item at an estimated proficiency level is normalized by the number of assertions in the item (similar to how information is computed for item sets in ELA and mathematics assessments). Details for CAI's adaptive testing algorithm are described in Appendix 2-N, Adaptive Algorithm Design.

A non-segmented test design was used for the MSA, meaning students received items from different disciplines in a random order (in contrast to a segmented test design, in which separate parts of the test are administered in a fixed order). In an adaptive test, the use of a non-segmented test design provides more freedom when selecting items targeting a current best estimate of proficiency. Embedded field-test items were randomly positioned in the test and randomly distributed across students. Every student received either one item cluster or four stand-alone items as field-test items throughout the test.

4.2 TEST BLUEPRINTS

Test blueprints provide the following guidelines:

- Length of the test
- Science disciplines to be covered and the acceptable number of items across PEs within each science discipline and Disciplinary Core Idea (DCI)

The test blueprints are presented in Table 12 - Table 13.

Table 12. Science Test Blueprint, Grade 5

Grade 5	Min Item Clusters	Max Item Clusters	Min Stand-Alone Items	Max Stand-Alone Items	Min Item Clusters + Min Stand-Alone Items	Max Item Clusters + Max Stand-Alone Items
Discipline—Physical Sciences, PE Total = 17	**2**	**2**	**4**	**4**	**6**	**6**
DCI—Motion and Stability: Forces and Interactions	**0**	**1**	**0**	**2**	**0**	**3**
3-PS2-1: Forces and Motion, Types of Interactions	0	1	0	1	0	1
3-PS2-2: Forces and Motion	0	1	0	1	0	1
3-PS2-3: Types of Interactions	0	1	0	1	0	1
3-PS2-4: Types of Interactions*	0	1	0	1	0	1
5-PS2-1: Types of Interactions, Space Systems	0	1	0	1	0	1
DCI—Energy	**0**	**1**	**0**	**2**	**0**	**3**
4-PS3-1: Energy	0	1	0	1	0	1
4-PS3-2: Conservation and Transfer of Energy	0	1	0	1	0	1
4-PS3-3: Conservation and Transfer of Energy, Energy and Forces	0	1	0	1	0	1
4-PS3-4: Conservation and Transfer of Energy*	0	1	0	1	0	1
5-PS3-1: Matter & Energy	0	1	0	1	0	1
DCI—Waves and Their Applications in Technologies for Information Transfer	**0**	**1**	**0**	**2**	**0**	**3**
4-PS4-1: Wave Properties	0	1	0	1	0	1
4-PS4-2: Electromagnetic Radiation	0	1	0	1	0	1
4-PS4-3: Information Technologies*	0	1	0	1	0	1
DCI—Matter and Its Interactions	**0**	**1**	**0**	**2**	**0**	**3**
5-PS1-1: Structure & Properties of Matter	0	1	0	1	0	1
5-PS1-2: Structure & Properties of Matter, Chemical Reactions	0	1	0	1	0	1
5-PS1-3: Structure & Properties of Matter	0	1	0	1	0	1
5-PS1-4: Chemical Reactions	0	1	0	1	0	1

Grade 5	Min Item Clusters	Max Item Clusters	Min Stand-Alone Items	Max Stand-Alone Items	Min Item Clusters + Min Stand-Alone Items	Max Item Clusters + Max Stand-Alone Items
Discipline—Life Sciences, PE Total = 12	**2**	**2**	**4**	**4**	**6**	**6**
DCI—From Molecules to Organisms: Structure and Function	**0**	**1**	**0**	**2**	**0**	**3**
3-LS1-1: Growth and Development of Organisms	0	1	0	1	0	1
4-LS1-1: Structure, Function	0	1	0	1	0	1
4-LS1-2: Information Processing	0	1	0	1	0	1
5-LS1-1: Matter & Energy Flow in Organisms	0	1	0	1	0	1
DCI—Ecosystems: Interactions, Energy, and Dynamics	**0**	**1**	**0**	**2**	**0**	**3**
3-LS2-1: Social Interactions and Group Behavior	0	1	0	1	0	1
5-LS2-1: Matter & Energy in Ecosystems	0	1	0	1	0	1
DCI—Inheritance and Variation of Traits	**0**	**1**	**0**	**2**	**0**	**3**
3-LS3-1: Inheritance and Variation of Traits	0	1	0	1	0	1
3-LS3-2: Inheritance and Variation of Traits	0	1	0	1	0	1
DCI—Biological Evolution: Unity and Diversity	**0**	**1**	**0**	**2**	**0**	**3**
3-LS4-1: Evidence of Common Ancestry and Diversity	0	1	0	1	0	1
3-LS4-2: Natural Selection	0	1	0	1	0	1
3-LS4-3: Adaptation	0	1	0	1	0	1
3-LS4-4: Ecosystems*	0	1	0	1	0	1
Discipline—Earth and Space Sciences, PE Total = 13	**2**	**2**	**4**	**4**	**6**	**6**
DCI—Earth's Systems	**0**	**1**	**0**	**2**	**0**	**3**
3-ESS2-1: Weather and Climate	0	1	0	1	0	1
3-ESS2-2: Weather and Climate	0	1	0	1	0	1
4-ESS2-1: Earth Materials and Systems	0	1	0	1	0	1

Grade 5	Min Item Clusters	Max Item Clusters	Min Stand-Alone Items	Max Stand-Alone Items	Min Item Clusters + Min Stand-Alone Items	Max Item Clusters + Max Stand-Alone Items
4-ESS2-2: Plate Tectonics and System Interactions	0	1	0	1	0	1
5-ESS2-1: Earth Materials and Systems	0	1	0	1	0	1
5-ESS2-2: Roles of Water in Earth's Surface Processes	0	1	0	1	0	1
DCI—Earth and Human Activity	**0**	**1**	**0**	**2**	**0**	**3**
3-ESS3-1: Natural Hazards*	0	1	0	1	0	1
4-ESS3-2: Natural Hazards*	0	1	0	1	0	1
4-ESS3-1: Natural Resources	0	1	0	1	0	1
5-ESS3-1: Human Impacts on Earth Systems	0	1	0	1	0	1
DCI – Earth's Place in the Universe	**0**	**1**	**0**	**2**	**0**	**3**
4-ESS1-1: Earth's History	0	1	0	1	0	1
5-ESS1-1: The Universe and Its Stars	0	1	0	1	0	1
5-ESS1-2: Earth and the Solar System	0	1	0	1	0	1
PE Total = 42	**6**	**6**	**12**	**12**	**18**	**18**

*These PEs have an engineering component. Bolded PEs indicate Montana-adapted standards.

Table 13. Science Test Blueprint, Grade 8

Grade 8	Min Item Clusters	Max Item Clusters	Min Stand-Alone Items	Max Stand-Alone Items	Min Item Clusters + Min Stand-Alone Items	Max Item Clusters + Max Stand-Alone Items
Discipline—Physical Sciences, PE Total = 18	**2**	**2**	**4**	**4**	**6**	**6**
DCI—Matter and Its Interactions	**0**	**1**	**0**	**2**	**0**	**3**
MS-PS1-1: Structure & Properties of Matter	0	1	0	1	0	1
MS-PS1-2: Structure & Properties of Matter, Chemical Reactions	0	1	0	1	0	1
MS-PS1-3: Structure & Properties of Matter, Chemical Reactions	0	1	0	1	0	1
MS-PS1-4: Structure & Properties of Matter, Energy	0	1	0	1	0	1
MS-PS1-5: Chemical Reactions	0	1	0	1	0	1
MS-PS1-6: Chemical Reactions*	0	1	0	1	0	1
DCI—Motion and Stability: Forces and Interactions	**0**	**1**	**0**	**2**	**0**	**3**
MS-PS2-1: Forces & Motion*	0	1	0	1	0	1
MS-PS2-2: Forces & Motion	0	1	0	1	0	1
MS-PS2-3: Types of Interactions	0	1	0	1	0	1
MS-PS2-4: Types of Interactions	0	1	0	1	0	1
MS-PS2-5: Types of Interactions	0	1	0	1	0	1
DCI—Energy	**0**	**1**	**0**	**2**	**0**	**3**
MS-PS3-1: Energy	0	1	0	1	0	1
MS-PS3-2: Energy, Relationship Between Energy and Forces	0	1	0	1	0	1
MS-PS3-3: Energy, Conservation of Energy*	0	1	0	1	0	1
MS-PS3-4: Energy, Conservation and Transfer of Energy	0	1	0	1	0	1
MS-PS3-5: Conservation and Transfer of Energy	0	1	0	1	0	1
DCI—Waves and Their Applications in Technologies for Information Transfer	**0**	**1**	**0**	**2**	**0**	**3**

Grade 8	Min Item Clusters	Max Item Clusters	Min Stand-Alone Items	Max Stand-Alone Items	Min Item Clusters + Min Stand-Alone Items	Max Item Clusters + Max Stand-Alone Items
MS-PS4-1: Wave Properties	0	1	0	1	0	1
MS-PS4-2: Wave Properties & Electromagnetic Radiation	0	1	0	1	0	1
Discipline—Life Sciences, PE Total = 19	**2**	**2**	**4**	**4**	**6**	**6**
DCI—From Molecules to Organisms: Structures and Processes	**0**	**1**	**0**	**2**	**0**	**3**
MS-LS1-1: Structure and Function	0	1	0	1	0	1
MS-LS1-2: Structure and Function	0	1	0	1	0	1
MS-LS1-3: Structure and Function	0	1	0	1	0	1
MS-LS1-4: Growth and Development of Organisms	0	1	0	1	0	1
MS-LS1-5: Growth and Development of Organisms	0	1	0	1	0	1
MS-LS1-6: Organization of Matter and Energy Flow in Organisms	0	1	0	1	0	1
MS-LS1-7: Organization of Matter and Energy Flow in Organisms	0	1	0	1	0	1
DCI—Ecosystems: Interactions, Energy, and Dynamics	**0**	**1**	**0**	**2**	**0**	**3**
MS-LS2-1: Interdependent Relationships in Ecosystems	0	1	0	1	0	1
MS-LS2-2: Interdependent Relationships in Ecosystems	0	1	0	1	0	1
MS-LS2-3: Cycle of Matter & Energy in Ecosystems	0	1	0	1	0	1
MS-LS2-5: Ecosystem Dynamics, Biodiversity and Humans*	0	1	0	1	0	1
DCI—Heredity: Inheritance and Variation of Traits	**0**	**1**	**0**	**2**	**0**	**3**
MS-LS3-1: Inheritance and Variation of Traits	0	1	0	1	0	1
MS-LS3-2: Inheritance and Variation of Traits	0	1	0	1	0	1
DCI—Biological Evolution: Unity and Diversity	**0**	**1**	**0**	**2**	**0**	**3**

Grade 8	Min Item Clusters	Max Item Clusters	Min Stand-Alone Items	Max Stand-Alone Items	Min Item Clusters + Min Stand-Alone Items	Max Item Clusters + Max Stand-Alone Items
MS-LS4-1: Evidence of Common Ancestry and Diversity	0	1	0	1	0	1
MS-LS4-2: Evidence of Common Ancestry and Diversity	0	1	0	1	0	1
MS-LS4-3: Natural Selection & Adaptation	0	1	0	1	0	1
MS-LS4-4: Natural Selection	0	1	0	1	0	1
MS-LS4-5: Natural Selection	0	1	0	1	0	1
MS-LS4-6: Adaptation	0	1	0	1	0	1
Discipline—Earth and Space Sciences, PE Total = 15	**2**	**2**	**4**	**4**	**6**	**6**
DCI—Earth's Place in the Universe	**0**	**1**	**0**	**2**	**0**	**3**
MS-ESS1-1: The Universe and Its Stars, Earth and the Solar System	0	1	0	1	0	1
MS-ESS1-2: The Universe and Its Stars, Earth and the Solar System	0	1	0	1	0	1
MS-ESS1-3: Earth and the Solar System	0	1	0	1	0	1
MS-ESS1-4: History of Earth	0	1	0	1	0	1
DCI—Earth's Systems	**0**	**1**	**0**	**2**	**0**	**3**
MS-ESS2-1: Earth's Materials & Systems	0	1	0	1	0	1
MS-ESS2-2: Earth's Materials & Systems, Roles of Water	0	1	0	1	0	1
MS-ESS2-3: Plate Tectonics	0	1	0	1	0	1
MS-ESS2-4: Roles of Water in Earth's Surface Processes	0	1	0	1	0	1
MS-ESS2-5: Roles of Water, Weather & Climate	0	1	0	1	0	1
MS-ESS2-6: Roles of Water, Weather & Climate	0	1	0	1	0	1
DCI—Earth and Human Activity	**0**	**1**	**0**	**2**	**0**	**3**
MS-ESS3-1: Natural Resources	0	1	0	1	0	1

Grade 8	Min Item Clusters	Max Item Clusters	Min Stand-Alone Items	Max Stand-Alone Items	Min Item Clusters + Min Stand-Alone Items	Max Item Clusters + Max Stand-Alone Items
MS-ESS3-2: Natural Hazards	0	1	0	1	0	1
MS-ESS3-3: Human Impacts*	0	1	0	1	0	1
MS-ESS3-4: Human Impacts	0	1	0	1	0	1
ME-ESS3-5: Global Climate Change	0	1	0	1	0	1
PE Total = 52	**6**	**6**	**12**	**12**	**18**	**18**

*These PEs have an engineering component. Bolded PEs indicate Montana-adapted standards.

The main characteristics of the blueprint were that any PE could be tested only once (indicated by the values of 0 and 1 for the minimum and maximum values of the individual PEs in Table 12 and Table 13). In general, no more than one item cluster or two stand-alone items could be sampled from the same DCI, and no more than three total items could be sampled from the same DCI (as indicated by the minimum and maximum values in the rows representing DCIs).

While tests are not timed, the Montana Office of Public Instruction (OPI) published estimated testing times for the Montana Science Assessment. The 85th percentile of the testing times (in minutes) is presented in Table 14.

Table 14. Montana Science Assessment 85th Percentile Testing Times by Grade

Subject	Grade	85th Percentile Testing
Science	5	95.15
	8	80.70

4.3 ONLINE TEST CONSTRUCTION

During fall 2022, CAI psychometricians and content experts worked with OPI content specialists and leadership to build item pools for the spring 2023 administration. The Montana Science Assessment test construction used a structured test construction plan, explicit blueprints, and active collaborative participation from all parties.

CAI test developers built the 2023 Montana Science Assessment test item pools to match items exactly to the detailed test blueprints. Operational items were selected from ten item banks (ICCR, Connecticut, Hawaii, Idaho, Montana, Rhode Island, Oregon, Utah, West Virginia, and Wyoming) to fulfill the blueprint for each grade. Table 15 through Table 19 summarize the 2023 Montana Science Assessment item pool. Appendix 2-M, Montana Science Assessment Item Pool, provides the 2023 item pool by grade band, PE, and origin.

Table 15. Spring 2023 Montana Science Assessment Operational and Field-Test Item Pool

Grade and Item Type	ICCR Items	Montana Items	MOU Items[a]	Total Pool Items
Grade 5	**102**	**42**	**81**	**225**
Cluster	36	13	42	91
Stand-Alone	66	29	39	134
Grade 8	**74**	**33**	**86**	**193**
Cluster	22	13	42	77
Stand-Alone	52	20	44	116
Total	**176**	**75**	**167**	**418**

[a]Other MOU states include Connecticut, Hawaii, Idaho, Oregon, Rhode Island, Utah, West Virginia, and Wyoming.

Table 16. Spring 2023 Montana Assessment Operational Item Pool

Grade and Item Type	ICCR Operational Items	Montana Operational Items	MOU Operational Items[a]	Total Operational Pool Items
Grade 5	**102**	**30**	**81**	**213**
Cluster	36	9	42	87
Stand-Alone	66	21	39	126
Grade 8	**74**	**24**	**86**	**184**
Cluster	22	8	42	72
Stand-Alone	52	16	44	112
Total	**176**	**54**	**167**	**397**

[a]Other MOU states include Connecticut, Hawaii, Idaho, Oregon, Rhode Island, Utah, West Virginia, and Wyoming.

Table 17. Spring 2023 Montana Science Assessment Field-Test Item Pool

Grade and Item Type	ICCR Field-Test Items	Montana Field-Test Items	MOU Field-Test Items[a]	Total Field-Test Pool Items
Grade 5	**0**	**12**	**0**	**12**
Cluster	0	4	0	4
Stand-Alone	0	8	0	8
Grade 8	**0**	**9**	**0**	**9**
Cluster	0	5	0	5
Stand-Alone	0	4	0	4
Total	**0**	**21**	**0**	**21**

[a]Montana only field tested Montana-owned items.

Table 18. Spring 2023 Montana Science Assessment Operational and Field-Test Item Pool by Science Discipline

Grade	Science Discipline	Item Type	ICCR Items	Montana Items	MOU Items[a]	Total Pool Items
5	Earth and Space Sciences	Cluster	15	5	13	33
		Stand-Alone	19	8	17	44
	Life Sciences	Cluster	10	4	10	24
		Stand-Alone	22	10	8	40
		Cluster	11	4	19	34

Grade	Science Discipline	Item Type	ICCR Items	Montana Items	MOU Items[a]	Total Pool Items
	Physical Sciences	Stand-Alone	25	11	14	50
8	Earth and Space Sciences	Cluster	8	4	12	24
		Stand-Alone	16	9	12	37
	Life Sciences	Cluster	7	5	11	23
		Stand-Alone	21	5	13	39
	Physical Sciences	Cluster	7	4	19	30
		Stand-Alone	15	6	19	40
Total			**176**	**75**	**167**	**418**

[a]Other MOU states include Connecticut, Hawaii, Idaho, Oregon, Rhode Island, Utah, West Virginia, and Wyoming.

Table 19. Spring 2023 Montana Science Assessment Operational and Field-Test Item Pool by Disciplinary Core Idea

Grade	Science Discipline	Disciplinary Core Idea	ICCR Items	Montana Items	MOU Items[a]	Total Pool Items
5	Earth and Space Sciences	ESS1	9	5	10	24
		ESS2	11	4	11	26
		ESS3	14	4	9	27
	Life Sciences	LS1	10	5	5	20
		LS2	5	2	9	16
		LS3	5	2	2	9
		LS4	12	5	2	19
	Physical Sciences	PS1	10	2	8	20
		PS2	10	7	11	28
		PS3	11	3	10	24
		PS4	5	3	4	12
8	Earth and Space Sciences	ESS1	6	4	6	16
		ESS2	8	4	8	20
		ESS3	10	5	10	25
	Life Sciences	LS1	9	2	9	20
		LS2	7	5	6	18
		LS3	2	2	5	9
		LS4	10	1	4	15
	Physical Sciences	PS1	6	5	18	29
		PS2	4	3	11	18
		PS3	8	1	7	16
		PS4	4	1	2	7
Total			**176**	**75**	**167**	**418**

[a]Other MOU states include Connecticut, Hawaii, Idaho, Oregon, Rhode Island, Utah, West Virginia, and Wyoming.

More information about *p*-values, biserial correlations, and IRT parameters can be found in Volume 1, Annual Technical Report. The details on calibration and scoring of the Montana Science Assessment can also be found in Volume 1 of this report.

4.4 BRAILLE ACCOMMODATION FORMS

Student scores should not depend on the mode of administration or type of test form. The Montana Science Assessment was primarily administered in an online test delivery system in spring 2023. MSA offers a braille paper fixed form, a braille online fixed form, and computer adaptive form. No student took the paper-pencil test or the braille test in grades 5 or 8. Nonetheless, scores obtained via alternate modes of administration must be established as comparable to scores obtained through online testing. To achieve this, content specialists began with the online pool when building the paper-pencil/braille forms. Next, they removed any items that could not be rendered on paper or transcribed in braille. Finally, content specialists constructed fixed forms that adhered to the test blueprint. In spring 2023, the paper-pencil/braille forms met all blueprint requirements.

5. SIMULATION SUMMARY REPORT

This section describes the results of the simulated test administrations used to configure and evaluate the adequacy of the item-selection algorithm, which was used to administer the 2022–2023 Montana Assessment for grades 5 and 8. Simulations were conducted to configure the settings of the algorithm and to evaluate whether individual tests adhered to the test blueprint.

Some important settings included “Select Candidate Set 1” (cset1) and “Select Candidate Set 2” (cset2), which represent subsets of the item pool that were eligible for item selection. Refer to Appendix 2-N, Adaptive Algorithm Design, for more details of the current item-selection algorithm. In spring 2023, cset1 and cset2 values were set to 10 and 1. Psychometricians reviewed the simulation results and configured settings based on some key diagnostics, including the following:

- **Match-to-Test Blueprint.** This diagnostic determines whether the tests have the correct number of test items overall and the appropriate proportion by content categories at each level of the content hierarchy, as specified in the test blueprints for every science grade.
- **Item Exposure Rate.** This diagnostic evaluates the utility of item pools and identifies overexposed and underexposed items.

These diagnostics are interrelated. For example, if the test pool for a particular content category is limited (i.e., there are only a few test items available), achieving a 100% match to the blueprint for this content level will lead to a high item exposure rate, which means that a large number of students will see the same items. The software system that performs the simulation allows adjustments to the setting parameters to attain the best possible balance among these diagnostics. The simulation involves an iterative process that reviews initial results, adjusts the system parameters, runs new simulations, reviews the new results, and repeats the exercise until an optimal balance is achieved. The final setting would then be applied for the operational tests.

5.1 Factors Affecting Simulation Results

Several factors may influence simulation results for an adaptive test administration. These factors include the following:

- *The proportional relationship between the pool and the constraints to be met.* Proportionally distributed pools tend to make better use of the pool (i.e., more uniform item exposure) and make it easier to meet blueprint and other constraints. For example, if the specifications call for at least one item cluster per disciplinary core idea (DCI), but the pool has no item cluster for some DCIs, it may be impossible to meet this constraint.
- *The correlational structure between constraints.* It is easier to satisfy a constraint if there are instances of the constraint at all levels of another constraint. For example, if stand-alone items within a discipline are associated only with a specific DCI, it may be difficult to meet both the desired distribution of content and the desired distribution of item type.
- *Whether or not there is a strict maximum on a given constraint.* This means that the requirement must be met exactly in each test administration.

5.2 Results of Simulated Test Administrations: English

This section presents the simulation results for the online tests administered in English, which is the test taken by most students (>99.90%). Simulations were evaluated for all content areas using 5,000 simulated cases per grade.

5.2.1 Summary of Blueprint Match

The simulation results showed no blueprint violations at all content levels for both grades.

5.2.2 Item Exposure

The simulator output also reports the degree to which the constraints set forth in the blueprints may yield greater exposure of items to students. This is reported by examining the percentage of test administrations in which an item appears. For instance, in a fixed paper-pencil form, 100% of the items appear on 100% of the test administrations because every test taker takes the same form. In an adaptive test with a sufficiently large item pool, it is expected that most of the items would appear on a relatively small percentage of the test administrations only.

When this condition holds, it suggests that test administrations between students are more or less unique. Therefore, the item exposure rate was calculated for each item by dividing the total number of test administrations in which an item appears by the total number of tests administered. Then the distribution of the item exposure rate (r) is reported in eight bins. The bins are r=0% (unused), 0%<r≤1%, 1%<r≤5%, 5%<r≤20%, 20%<r≤40%, 40%<r≤60%, 60%<r≤80%, and 80%<r≤100%. If global item exposure is minimal, it is expected that the largest proportion of items would appear in the bins of 0%<r≤20%, which is an indication that most of the items appear on a very small percentage of the test forms.

Table 20 presents the percentage of items that fell into each exposure bin for all grades on the 2023 assessments. Most of the items had item exposure rates less than 20%. The minimum exposure rate was 1.34% in grade 5 and 0.46% in grade 8. No items had an exposure rate higher than 60% in grades 5 and 8. The item exposure rate for field test items ranged from 16%–18.1% across both grades. Table 20. Item Exposure Rates by Grade: Percentage of Items by Exposure Rate, Across All English Online Simulation Sessions

Grade	Total Items	[0,0]%	[0,1]%	[1,5]%	[5,20]%	[20,40]%	[40,60]%	[60,80]%	[80,100]%
5	213	0	0	30.04	64.79	3.76	1.51	0	0
8	184	0	8.70	36.96	40.76	12.50	10.87	0	0

5.3 Results of Simulated Test Administrations: Spanish

This section presents the simulation results for the tests administered in Spanish. The Spanish item pool consisted of a subset of ICCR items and some MOU items for which Spanish translations were available. Table 21 presents the number of items available for the Spanish tests in spring 2023.

Table 21. Spring 2023 Spanish Operational Item Pool

Grade	Item Type	Number of Items
5	Cluster	30
	Stand-Alone	43
8	Cluster	33
	Stand-Alone	45
Total		**151**

Simulations were evaluated for all content areas using 1,000 simulated cases per grade.

5.3.1 Summary of Blueprint Match

The simulation results showed no blueprint violations at all content levels for both grades.

5.3.2 Item Exposure

Table 22 presents the percentage of items that fell into each exposure bin for all grades. Most items were administered in more than 20% of the test administrations. Due to the limited size of the Spanish pool, most items had exposure rates of less than 40% rather than less than 20%. Four items had an exposure rate higher than 60% because of the limited Spanish item pool for some content categories.

Table 22. Item Exposure Rates by Grade: Percentage of Items by Exposure Rate, Across All Spanish Simulation Sessions

Grade	Total Items	[0,0]%	[0,1]%	[1,5]%	[5,20]%	[20,40]%	[40,60]%	[60,80]%	[80,100]%
5	73	0	0	0	49.10	47.95	9.59	1.37	0
8	78	0	0	11.54	38.46	33.33	12.82	3.85	0

6. OPERATIONAL TEST ADMINISTRATION SUMMARY REPORT

This section presents the blueprint match reports and item exposure rates for the spring 2023 operational test administrations.

6.1 BLUEPRINT MATCH

No blueprint violations were found at any content level for both grades, for both English and Spanish assessments.

6.2 ITEM EXPOSURE

Table 23 presents the item exposure rates for the spring 2023 test administration. The exposure rates closely resembled the simulation results described in Section 5.2.2, Item Exposure, for the English test administrations. More items on the Spanish tests had high exposure rates as compared with items on the English tests because of a smaller item pool. Also, the operational exposure rates differed from the simulation results because of small population sizes in both grades. In spring 2022, fewer than 20 students took the Spanish test in each grade.

Table 23. Item Exposure Rates by Grade: Percentage of Items by Exposure Rate, Across All Spring 2023 Test Administrations

Grade	Total Items	[0,0]%	[0,1]%	[1,5]%	[5,20]%	[20,40]%	[40,60]%	[60,80]%	[80,100]%
English									
5	213	0	0	30.05	65.26	3.29	1.41	0	0
8	184	0	17.39	22.83	46.2	11.96	1.63	0	0
Spanish									
5	60	0	0	0	30	48.33	16.67	5	0
8	62	0	0	0	45.16	29.03	11.29	9.68	4.84

7. References

Council of Chief State School Officers (CCSSO). (2015). *Science Assessment Item Collaborative (SAIC) assessment framework for the Next Generation Science Standards*. Washington, DC: Council of Chief State School Officers. Retrieved from https://ccsso.org/sites/default/files/2017-12/SAICAssessmentFramework_FINAL.pdf

National Research Council. (2012). *A framework for K–12 science education: Practices, crosscutting concepts, and core ideas*. Washington, DC: The National Academies Press.

Appendix 2-A

Item Writer Training Materials

Exhibit A-1. LABS Guidelines

LABS Guidelines

1. Stereotyping

Testing materials should not present persons stereotyped according to the following characteristics:

- Age
- Disability
- Gender
- Race/Ethnicity
- Sexual orientation

2. Sensitive or Controversial Subjects

Controversial or potentially distressing subjects should be avoided or treated sensitively. For example, a passage discussing the historical importance of a battle is acceptable whereas a graphic description of a battle would not be. Controversial subjects include:

- Death and Disease
- Gambling*
- Politics (Current)
- Race relations
- Religion
- Sexuality
- Superstition
- War

**References to gambling should be avoided in mathematics items related to probability.*

3. Advice

Testing materials should not advocate specific lifestyles or behaviors except in the most general or universally agreed-upon ways. For example, a recipe for a healthful fruit snack is acceptable but a passage recommending a specific diet is not. The following categories of advice should be avoided:

- Religion
- Sexual preference
- Exercise
- Diet

4. Dangerous Activity

Tests should not contain content that portrays people engaged in or explains how to engage in dangerous activities. Examples of dangerous activities include:

- Deep-sea diving
- Stunts
- Parachuting
- Smoking
- Drinking

5. Population Diversity and Ethnocentrism

Testing materials should:

- Reflect the diversity of the testing population
- Use stimulus materials (such as works of literature) produced by members of minority communities
- Use personal names from different ethnic origin communities
- Use pictures of people from different ethnic origin communities
- Avoid *ethnocentrism*, or the attitude that all people should share a particular group's language, beliefs, culture, or religion

6. Differential Familiarity and Elitism

Specialized concepts and terminology extraneous to the core content of test questions should be avoided. This caveat applies to terminology from the fields of:

- Construction
- Finance
- Sports
- Law
- Machinery
- Military topics
- Politics
- Science
- Technology
- Agriculture

7. Language Use

Language should be as inclusive as possible.

- Avoid masculine-coded words like mankind, manmade, and the generic "he"
- Use equal pairs such as husband and wife rather than man and wife

8. Language Accessibility

The grammar and vocabulary should be clear, concise, and appropriate for the intended grade level. The following should be avoided or used with care:

- Passive constructions
- Idioms
- Multiple subordinate clauses
- Pronouns with unclear antecedents
- Multiple-meaning words
- Non-standard grammar
- Dialect
- Jargon

9. Illustrations and Graphics

Illustrations and graphics should embody all of the previously referenced LABS Guidelines.

Exhibit A-2. LABS Checklist

LABS–Checklist

Stereotyping Considerations

- ☐ Does the material negatively represent, or stereotype people based on gender or sexual preference?
- ☐ Does the material portray one or more people with disabilities in a negative or stereotypical manner?
- ☐ Does the material portray one or more religious groups as aggressive or violent?
- ☐ Does the material romanticize or demean people based on socioeconomic status?
- ☐ Does the material portray one or more ethnic groups or cultures participating in certain stereotypical activities or occupations?
- ☐ Does the material portray one or more age groups in a negative or stereotypical manner?

Sensitive/Controversial Material Considerations

- ☐ Does the material require a student to take a position that challenges authority?
- ☐ Does the material present war or violence in an overly graphic manner?
- ☐ Does the material present sensitive or highly controversial subjects, such as death, war, abortion, euthanasia, or natural disasters, except where they are needed to meet State Content Standards?
- ☐ Does the material require test takers to disclose values that they would rather hold confidential?
- ☐ Does the material present sexual innuendoes?
- ☐ Does the material trivialize significant or tragic human experiences?
- ☐ Does the material require the parent, teacher, or test taker to support a position that is contrary to their religious beliefs?

ADVICE CONSIDERATIONS

- ☐ Does the material contain advice pertaining to health and well-being about which there is not a universal agreement?

POPULATION DIVERSITY

- ☐ Is the material written by members of diverse groups?
- ☐ Does the material reflect the experiences of diverse groups?
- ☐ Does the material portray people in positive nontraditional roles?
- ☐ Does test material represent the racial and ethnic composition of the testing population?
- ☐ Does the material reflect ethnocentrism?
- ☐ Does the material refer to population subgroups accurately?
- ☐ Does test material reflect diversity through the use of names, cultural references, pictures, and roles?

DIFFERENTIAL FAMILIARITY/ELITISM

- ☐ Does the material contain phrases, concepts, and beliefs that are irrelevant to testing domain and are likely to be more familiar to specific groups that others?
- ☐ Does the material require knowledge of individuals, events, or groups that is not familiar to all groups of students?
- ☐ Does the material suggest that affluence is related to merit or intelligence?
- ☐ Does the material suggest that poverty is related to increased negative behaviors in society?
- ☐ Does the material use language, content, or context that is offensive to people of a particular economic status?
- ☐ Does success with the material assume that the test taker has experience with a certain type of family structure?
- ☐ Does the material favor one socioeconomic group over another?
- ☐ Does the material assume values not shared by all test takers?

LINGUISTIC FEATURES/LANGUAGE ACCESSIBILITY/GRAPHICS

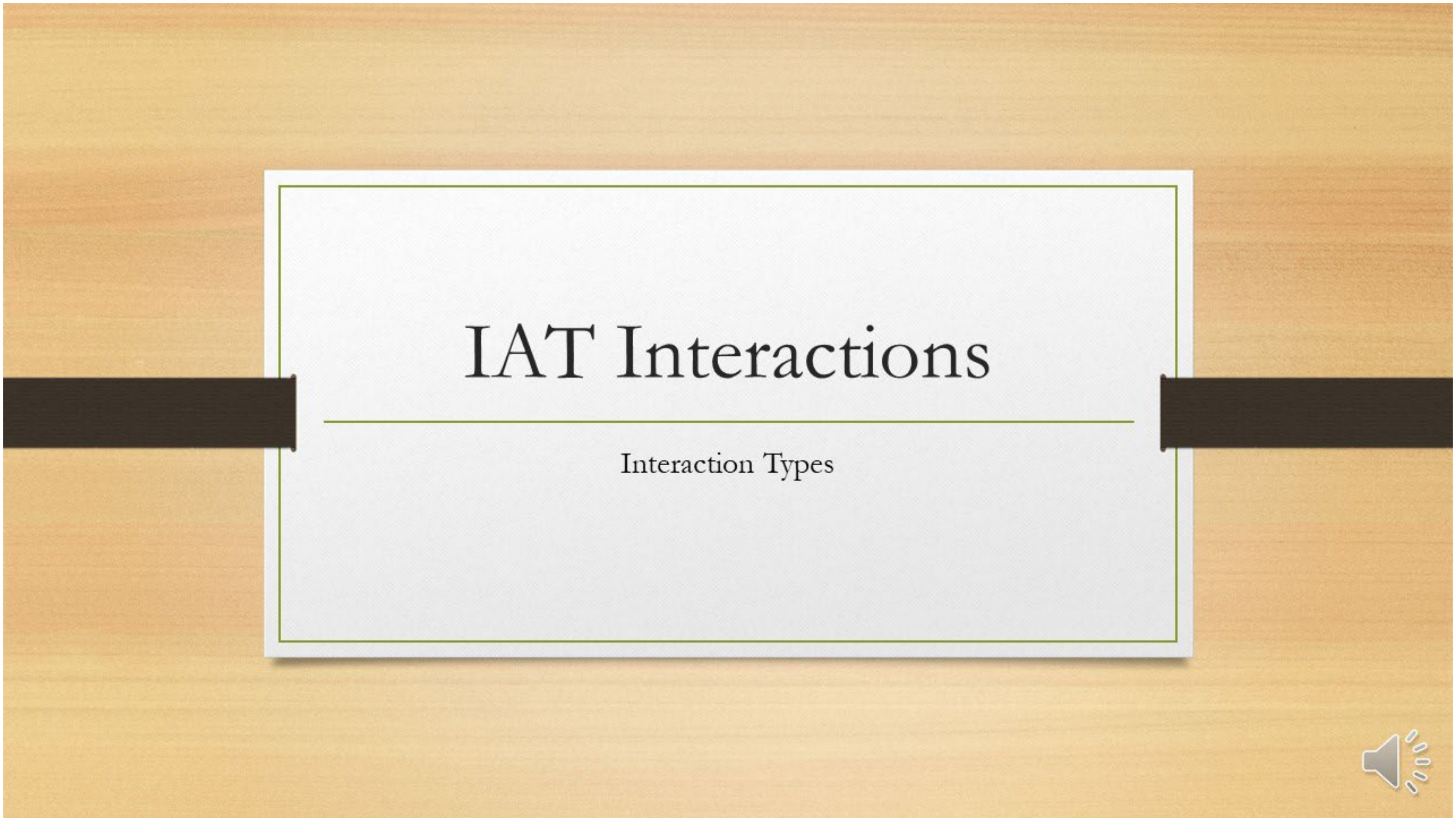
- ☐ Is grammar and vocabulary used in the items clear, concise, and appropriate for the intended grade level?
- ☐ Are passages at a difficulty level that is appropriate for the intended grade level?

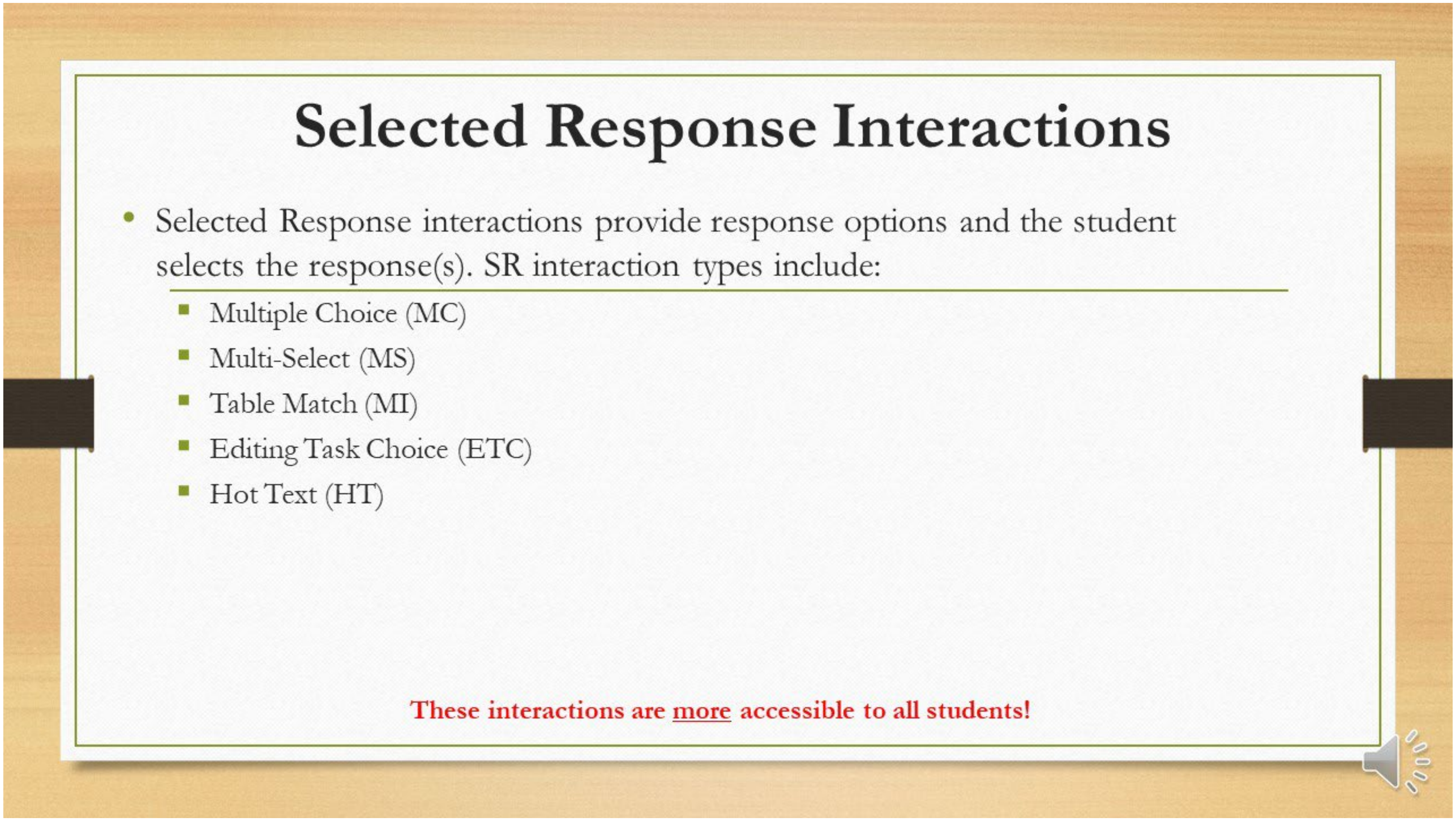
- Do the illustrations and graphics embody all of the previously referenced LABS Guidelines?

OTHER QUESTIONS TO CONSIDER

- Does the material favor one age group over others except in a context where experience or maturation is relevant?
- Does the material use language, content, or context that is not accessible to one or more of the age groups tested?
- Does the material contain language or content that contradicts values held by a certain culture?
- Does the material favor one racial or ethnic group over others?
- Does the material degrade people based on physical appearance or any physical, cognitive, or emotional challenge?
- Does the material focus only on a person's disability rather than portraying the whole person?
- Does the material favor one religion and/or demean others?

Exhibit A-3. An Overview of Interaction Types


IAT Interactions
Interaction Types


Selected Response Interactions
Selected Response interactions provide response options and the student selects the response(s). SR interaction types include:
Multiple Choice (MC)
Multi-Select (MS)
Table Match (MI)
Editing Task Choice (ETC)
Hot Text (HT)
These interactions are more accessible to all students!

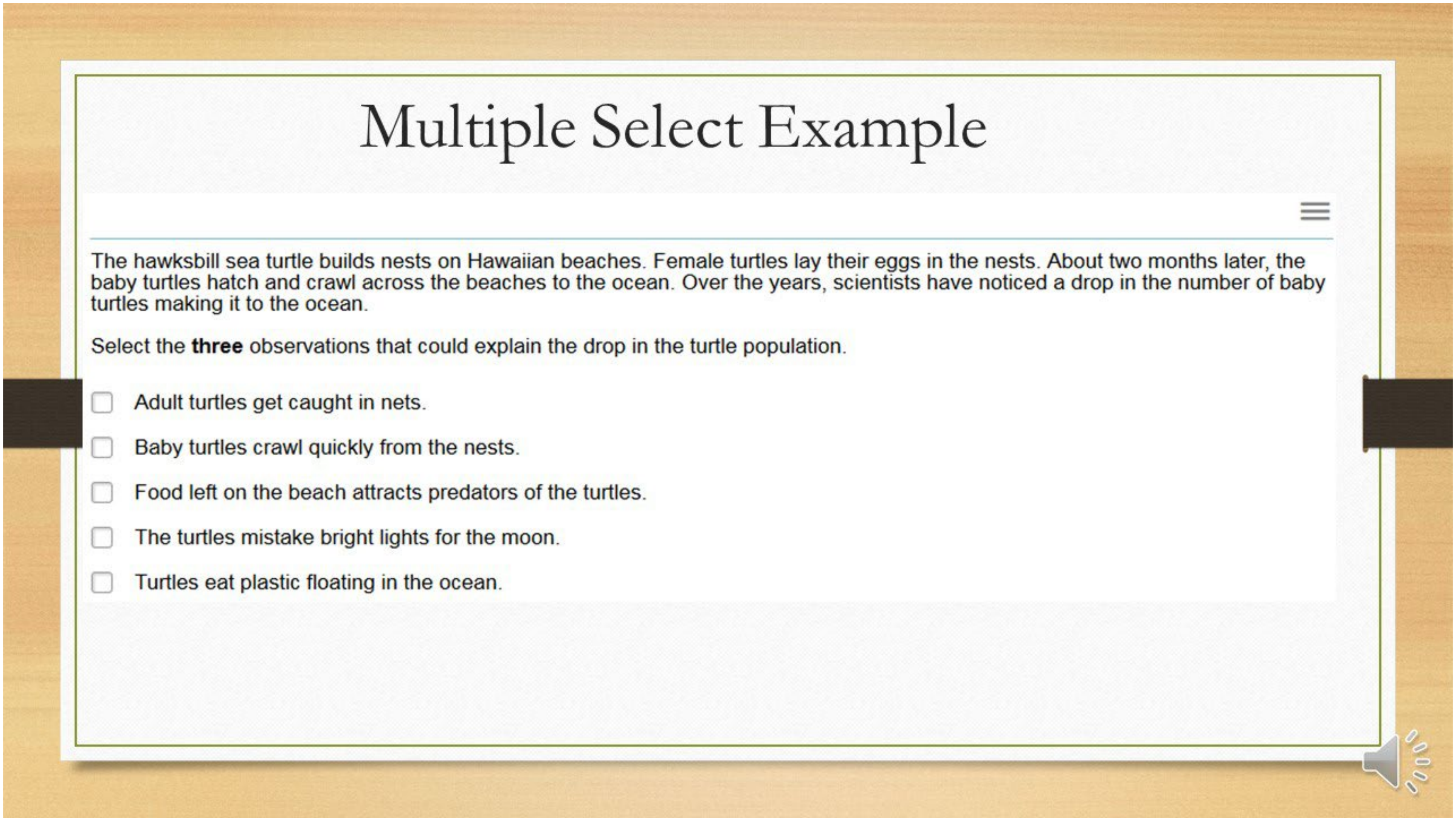
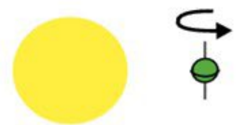
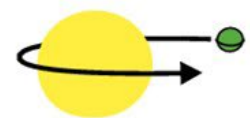

Multiple Select Example
The hawksbill sea turtle builds nests on Hawaiian beaches. Female turtles lay their eggs in the nests. About two months later, the baby turtles hatch and crawl across the beaches to the ocean. Over the years, scientists have noticed a drop in the number of baby turtles making it to the ocean.
Select the three observations that could explain the drop in the turtle population.
Adult turtles get caught in nets.
Baby turtles crawl quickly from the nests.
Food left on the beach attracts predators of the turtles.
The turtles mistake bright lights for the moon.
Turtles eat plastic floating in the ocean.

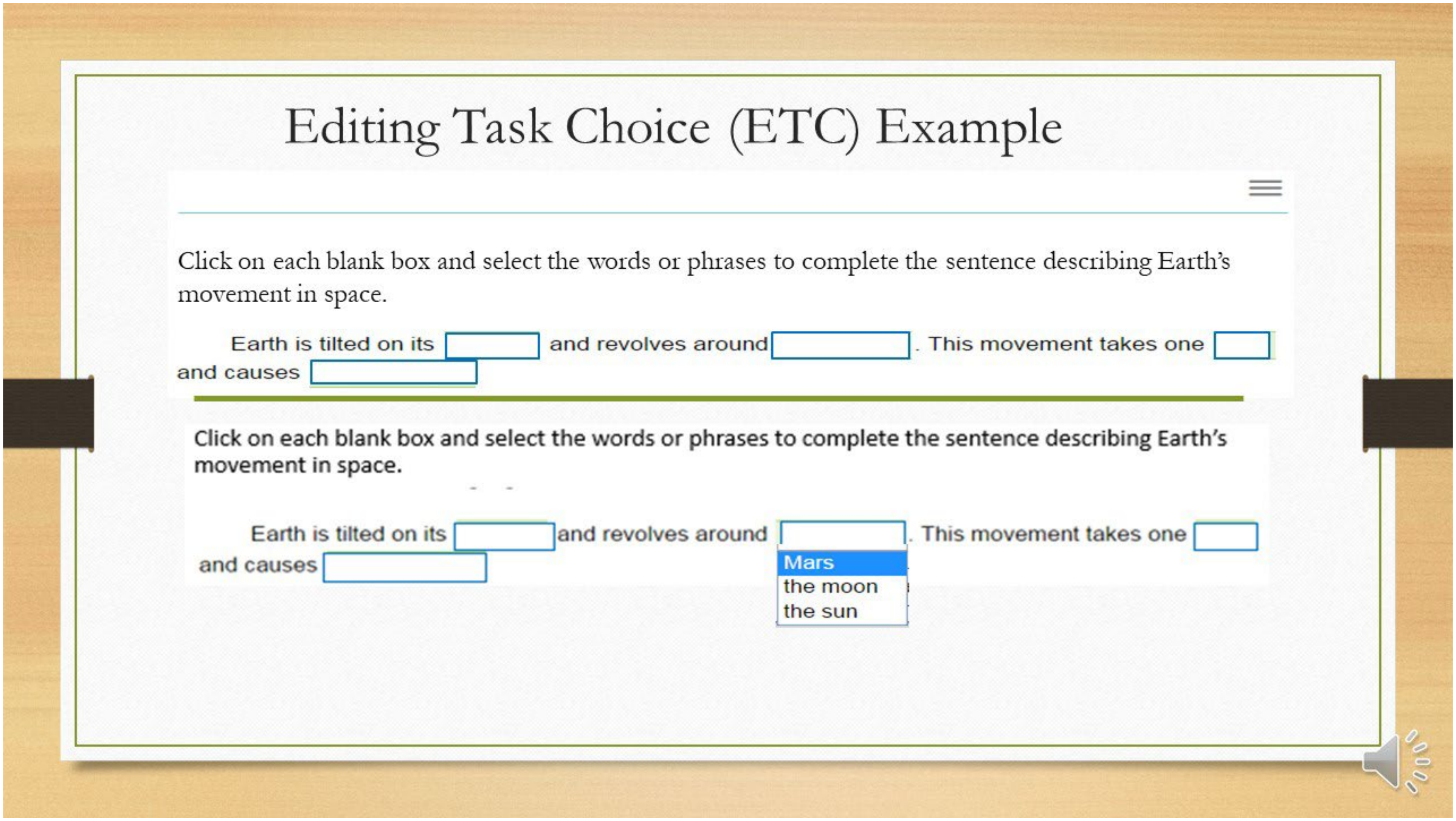
Table Match (MI) Example

Students use a large yellow ball and a small green ball to model the sun and Earth. They use the balls to explain the cause of day and night, to model the length of a year, and to explain the cause of the seasons.

Select **each** box to identify which movements of the balls are needed to explain each phenomenon.

- You can select more than one box for each statement.

	 Large yellow ball is stationary, while small green ball spins.	 Large yellow ball is stationary, while small green ball is tilted.	 Large yellow ball is stationary, while small green ball moves around it.
The cause of day and night	☐	☐	☐
The length of a year	☐	☐	☐
The cause of the seasons	☐	☐	☐


Editing Task Choice (ETC) Example
Click on each blank box and select the words or phrases to complete the sentence describing Earth's movement in space.
Earth is tilted on its and revolves around . This movement takes one and causes
Click on each blank box and select the words or phrases to complete the sentence describing Earth's movement in space.
Earth is tilted on its and revolves around . This movement takes one and causes
Mars
the moon
the sun

Hot Text (HT draggable) Example

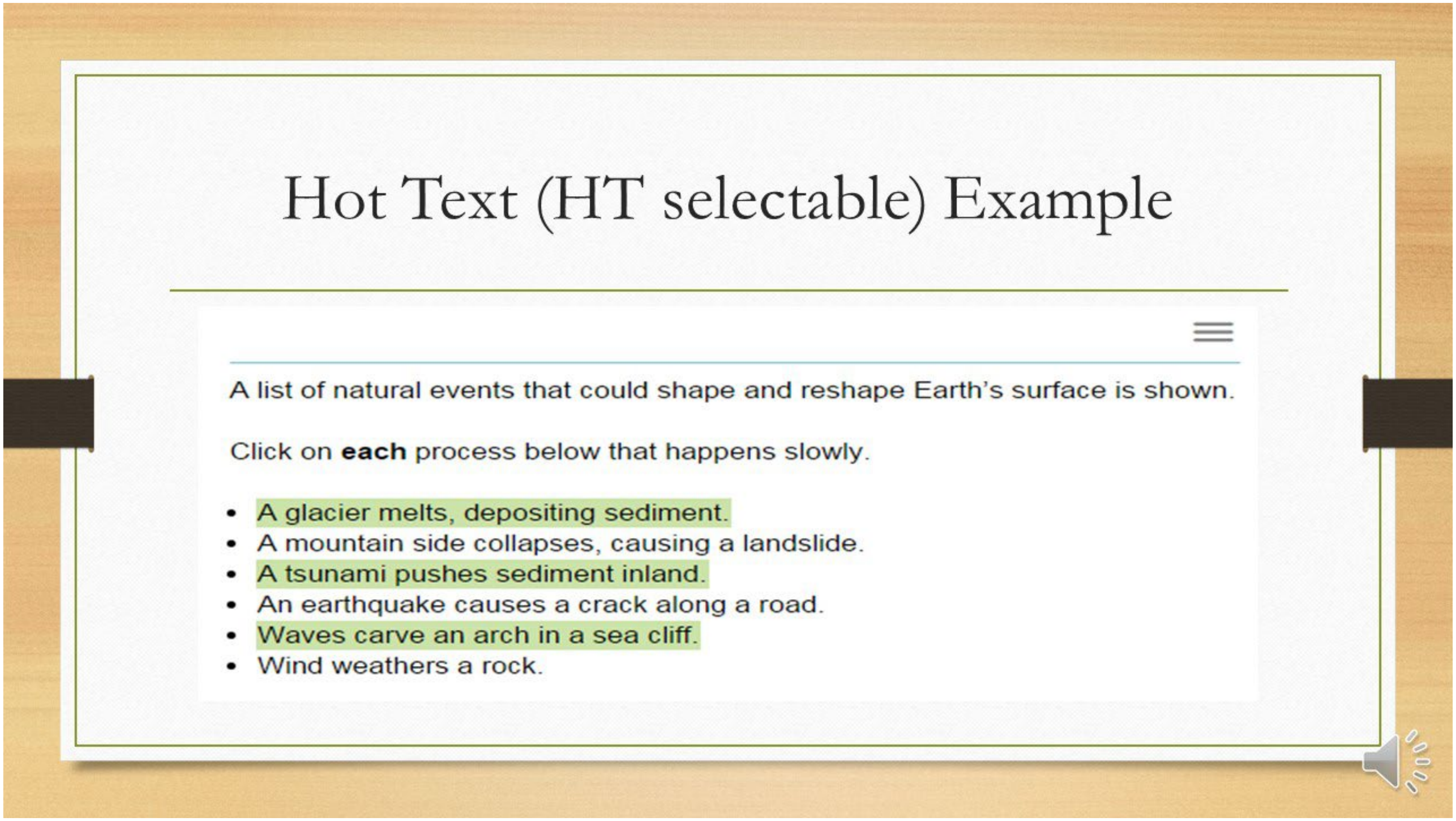
A list of natural events is shown.

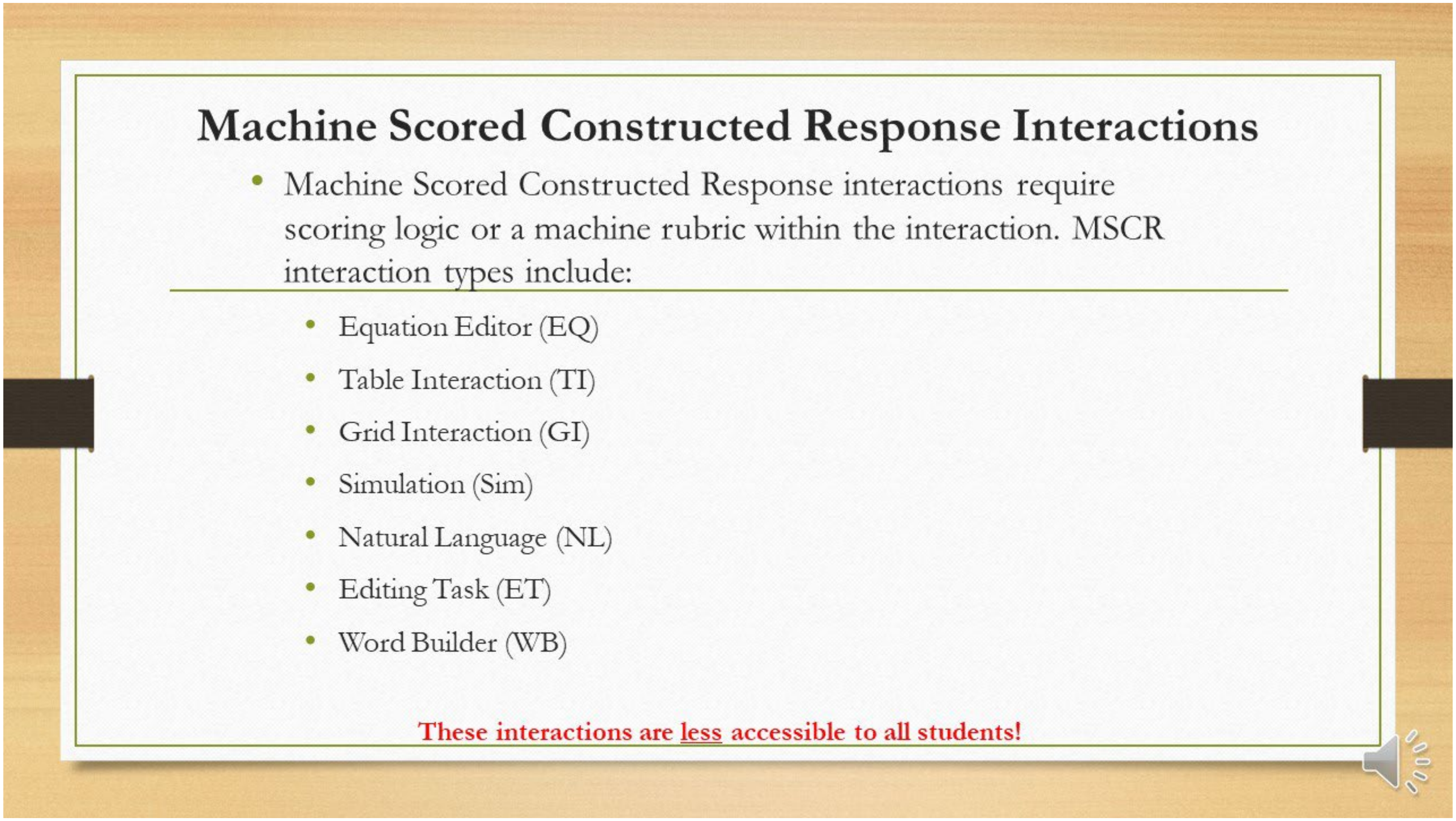
Click and drag the natural events to classify each natural event as either a fast or slow process that could shape and reshape Earth's surface.

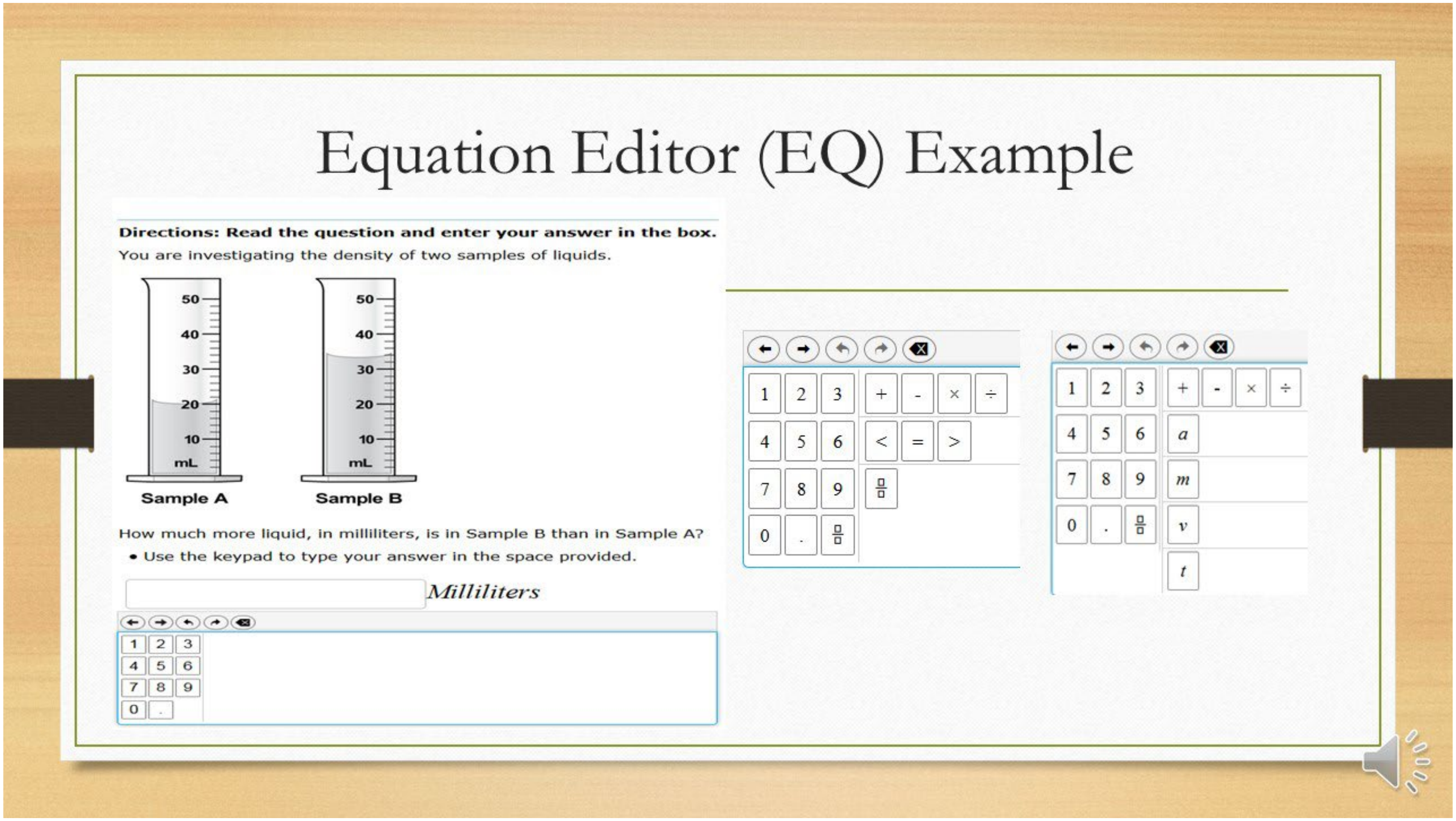
Fast and Slow Processes

Fast Process	Slow Process

1. A glacier melts, depositing sediment.
2. A mountain side collapses, causing a landslide.
3. A tsunami pushes sediment inland.
4. An earthquake causes a crack along a road.
5. Waves carve an arch in a sea cliff.
6. Wind weathers a rock.


Hot Text (HT selectable) Example
A list of natural events that could shape and reshape Earth's surface is shown.
Click on **each** process below that happens slowly.
• A glacier melts, depositing sediment.
• A mountain side collapses, causing a landslide.
• A tsunami pushes sediment inland.
• An earthquake causes a crack along a road.
• Waves carve an arch in a sea cliff.
• Wind weathers a rock.


Machine Scored Constructed Response Interactions
Machine Scored Constructed Response interactions require scoring logic or a machine rubric within the interaction. MSCR interaction types include:
Equation Editor (EQ)
Table Interaction (TI)
Grid Interaction (GI)
Simulation (Sim)
Natural Language (NL)
Editing Task (ET)
Word Builder (WB)
These interactions are less accessible to all students!


Equation Editor (EQ) Example
Directions: Read the question and enter your answer in the box.
You are investigating the density of two samples of liquids.
Sample A
Sample B
How much more liquid, in milliliters, is in Sample B than in Sample A?
• Use the keypad to type your answer in the space provided.
Milliliters

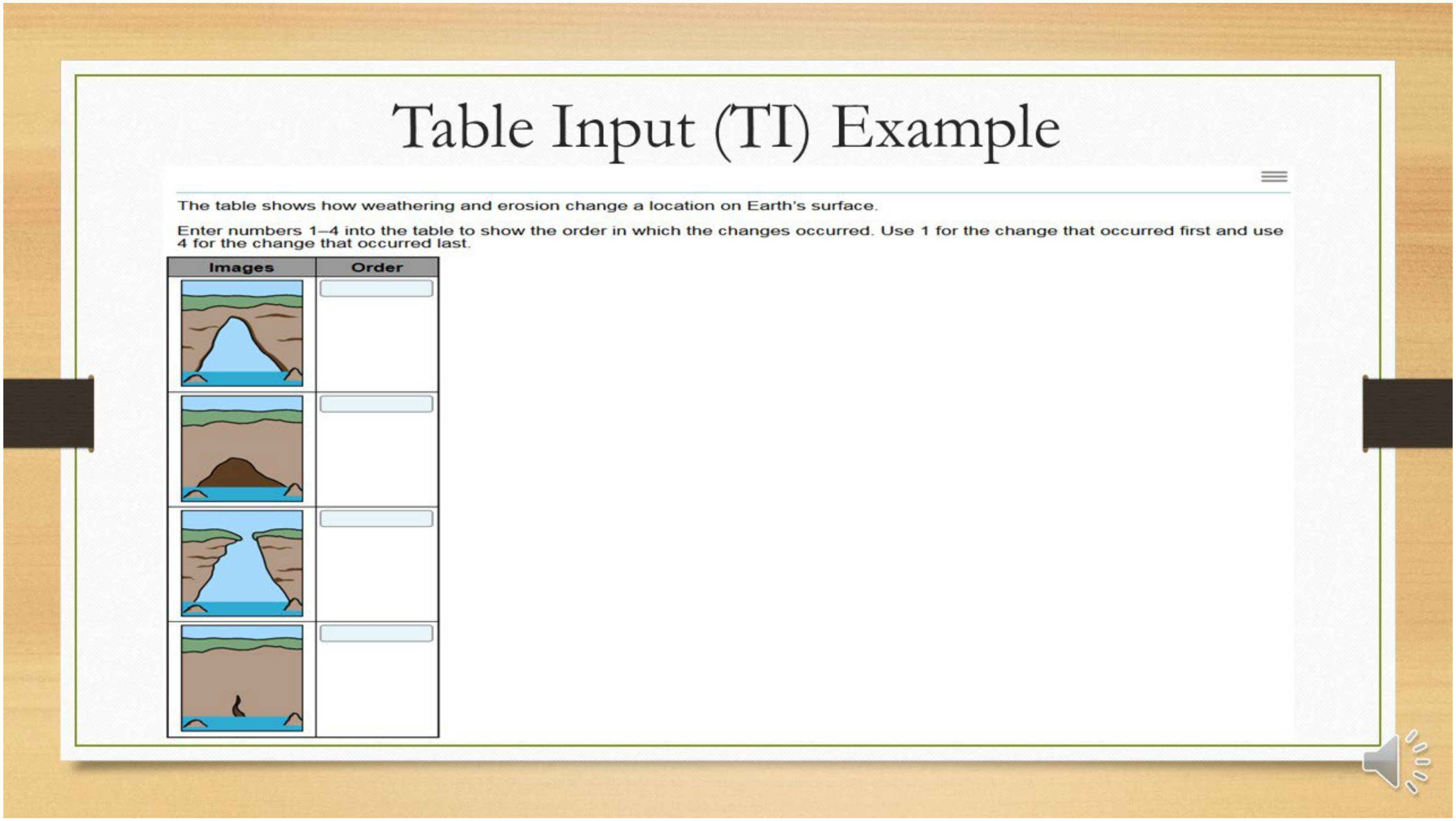
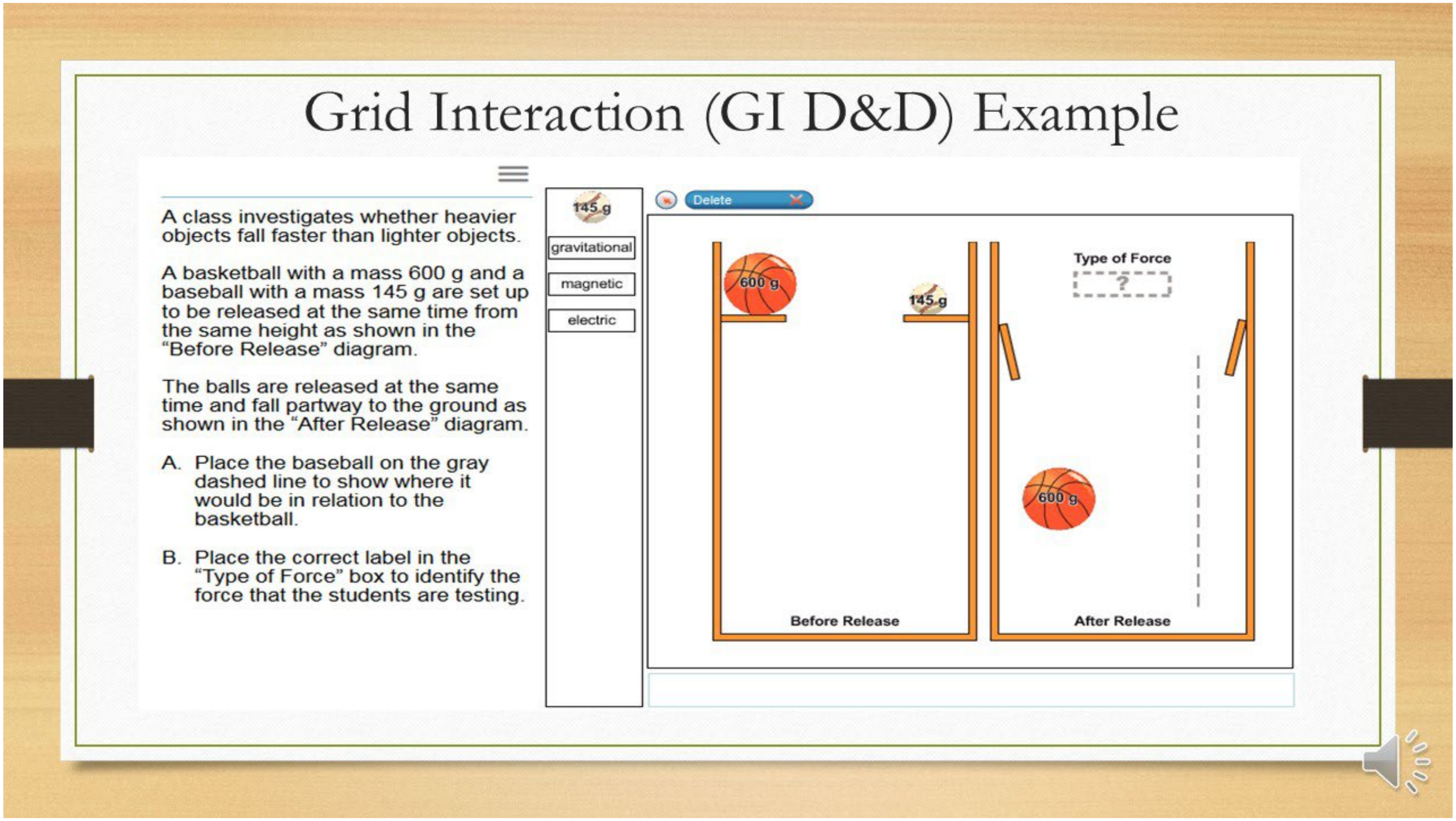
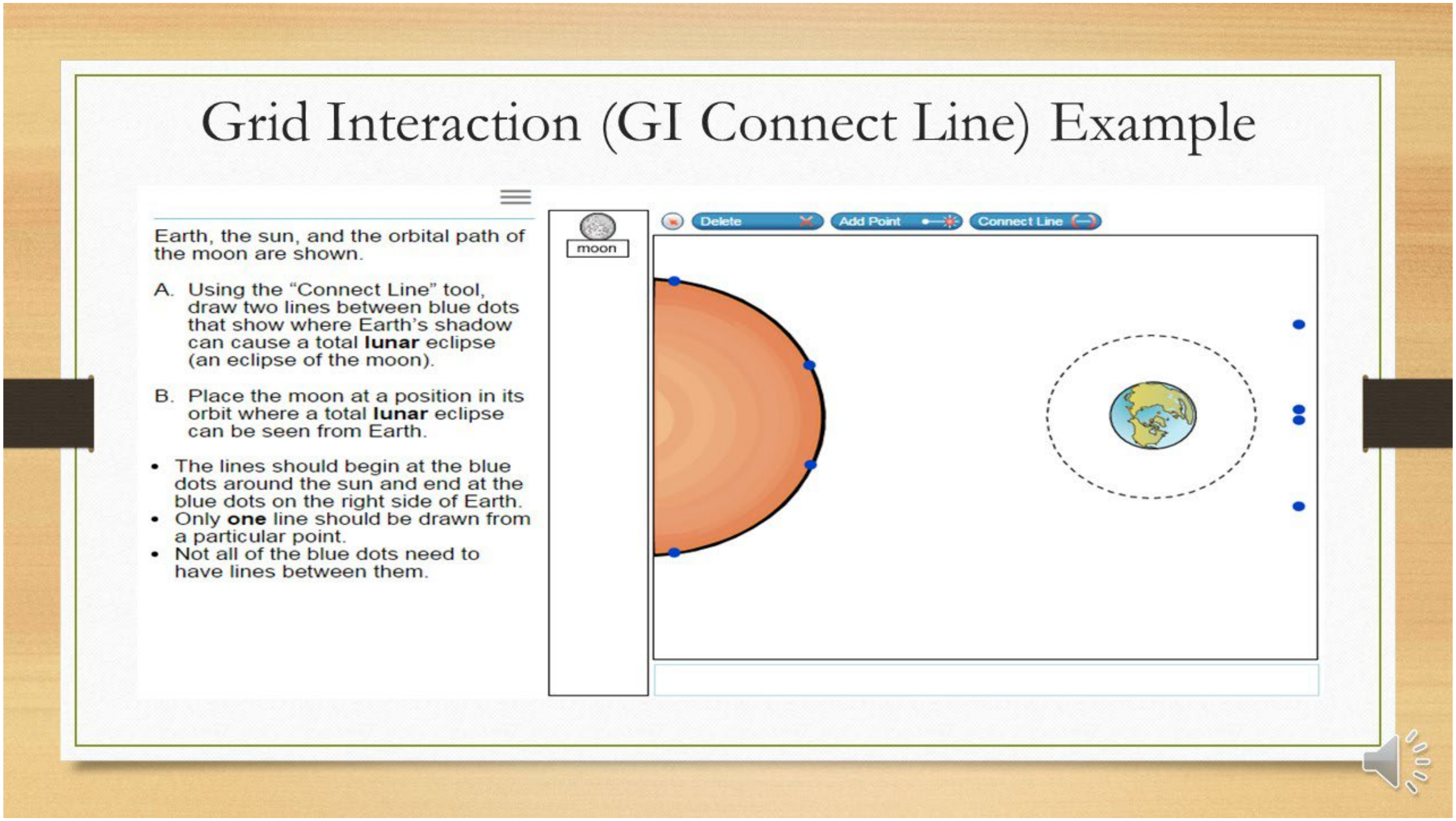
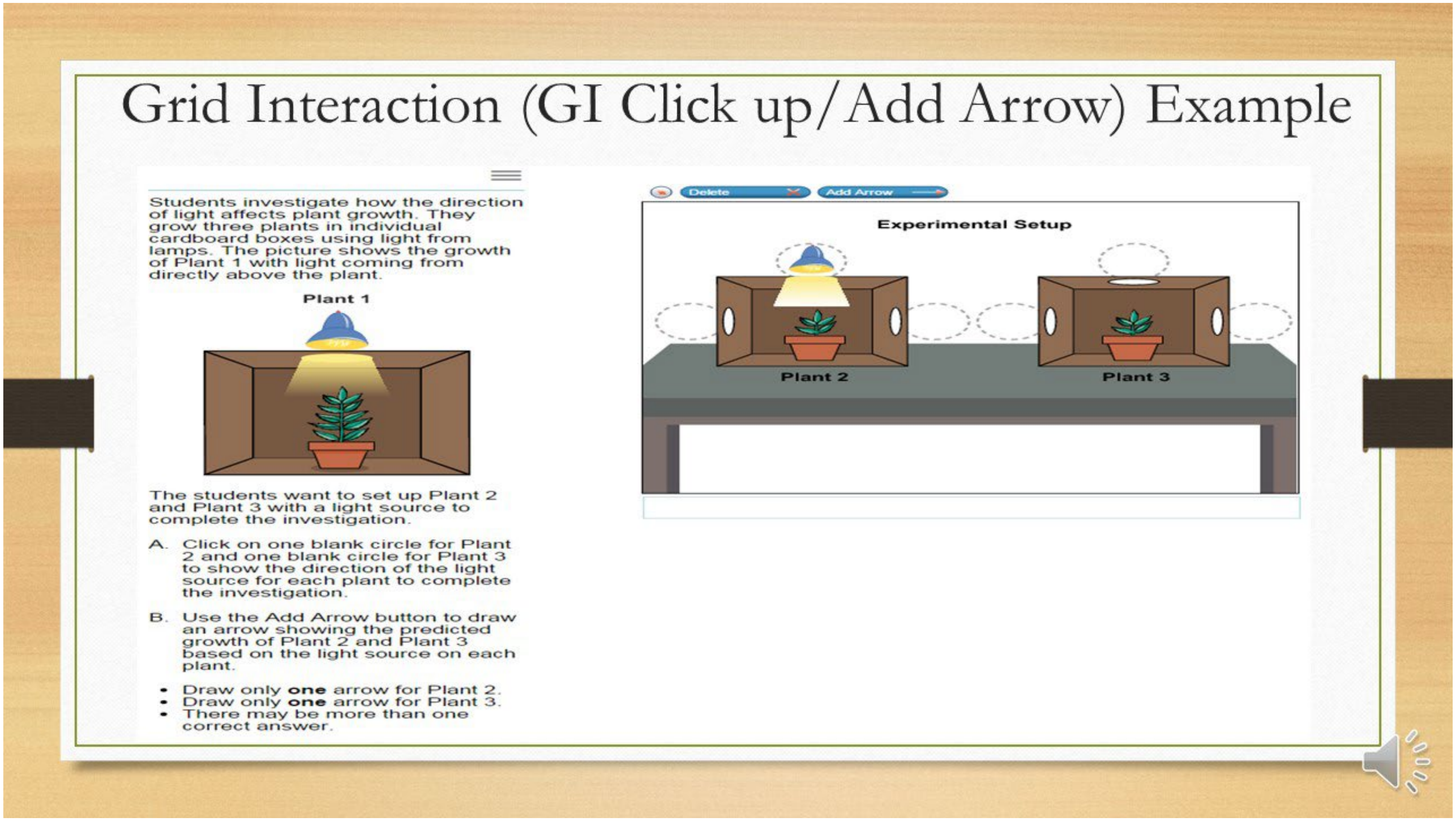
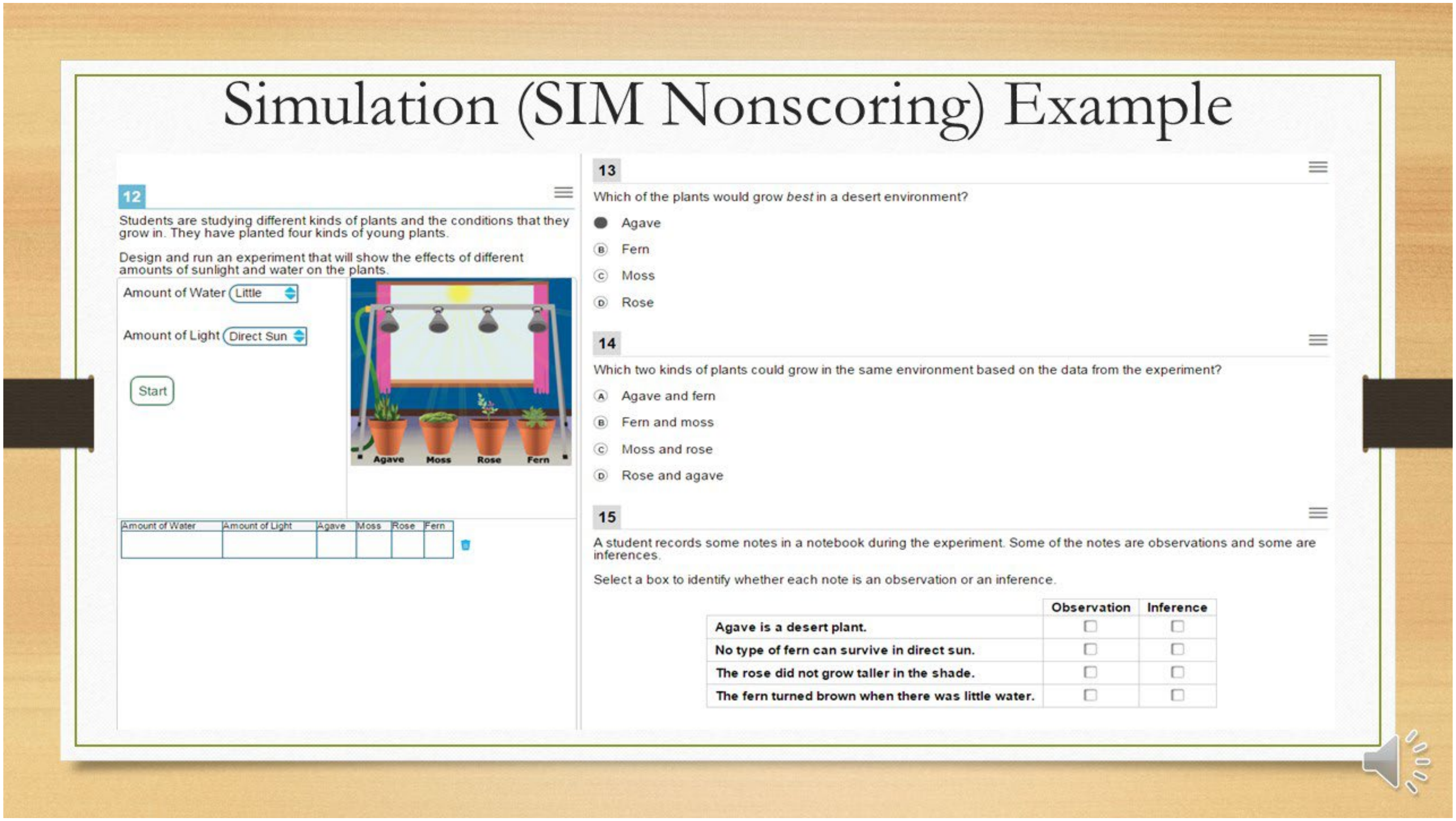


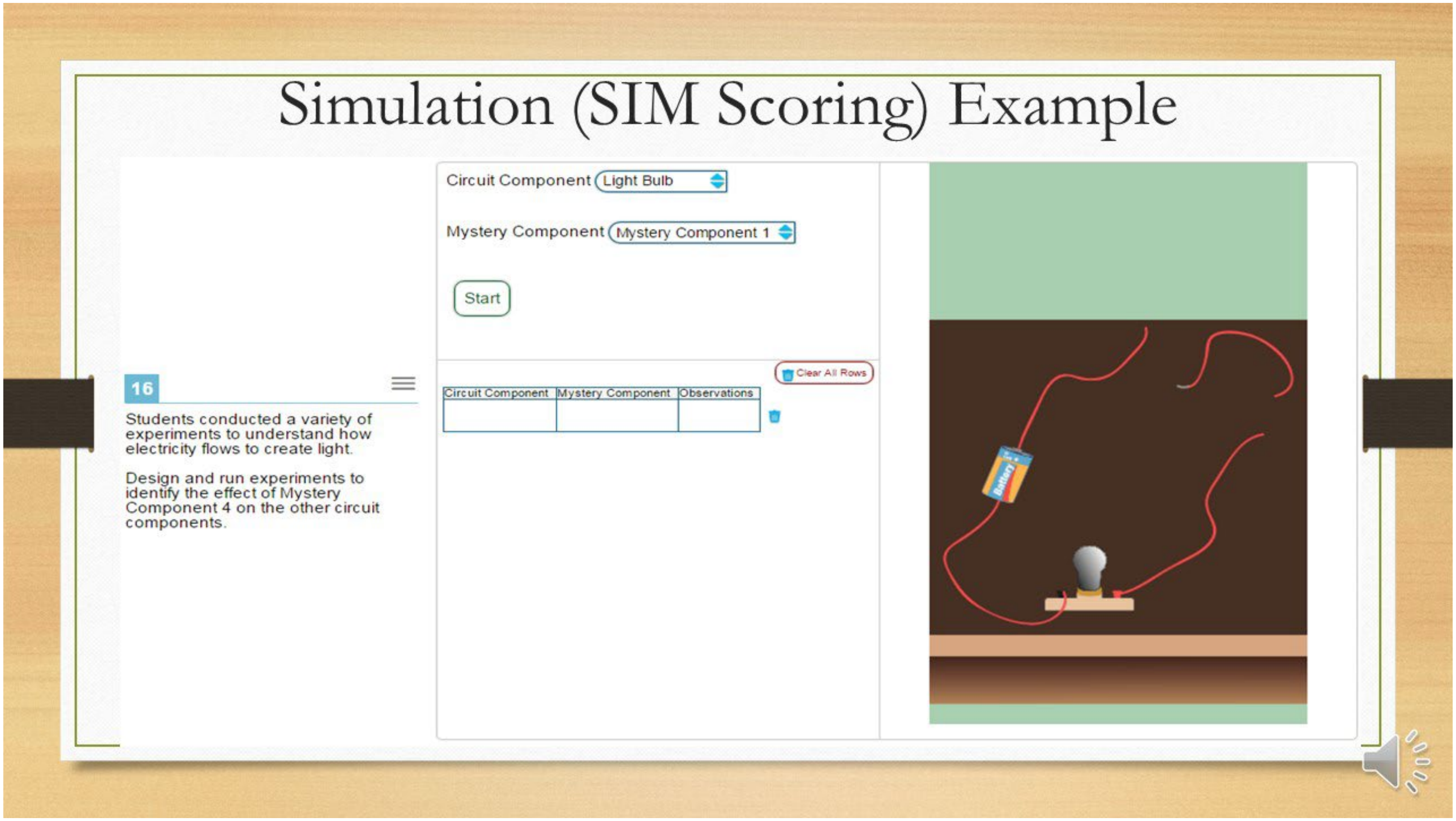
Table Input (TI) Example
The table shows how weathering and erosion change a location on Earth's surface.
Enter numbers 1–4 into the table to show the order in which the changes occurred. Use 1 for the change that occurred first and use 4 for the change that occurred last.
Images
Order

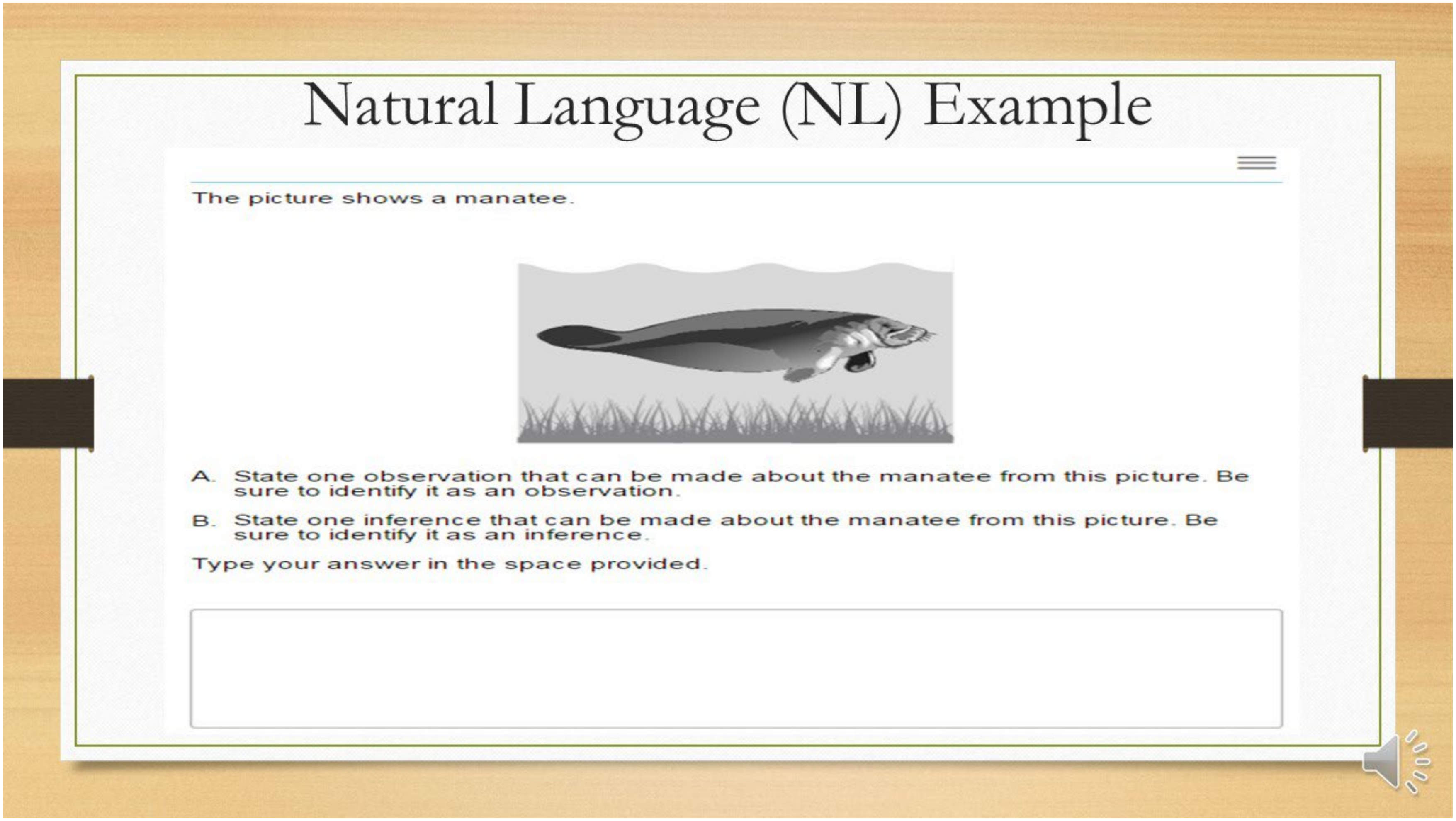

Grid Interaction (GI D&D) Example
A class investigates whether heavier objects fall faster than lighter objects.
A basketball with a mass 600 g and a baseball with a mass 145 g are set up to be released at the same time from the same height as shown in the "Before Release" diagram.
The balls are released at the same time and fall partway to the ground as shown in the "After Release" diagram.
A. Place the baseball on the gray dashed line to show where it would be in relation to the basketball.
B. Place the correct label in the "Type of Force" box to identify the force that the students are testing.
145 g
gravitational
magnetic
electric
Delete
600 g
145 g
Type of Force
?
600 g
Before Release
After Release


Grid Interaction (GI Connect Line) Example
Earth, the sun, and the orbital path of the moon are shown.
A. Using the "Connect Line" tool, draw two lines between blue dots that show where Earth's shadow can cause a total **lunar** eclipse (an eclipse of the moon).
B. Place the moon at a position in its orbit where a total **lunar** eclipse can be seen from Earth.
• The lines should begin at the blue dots around the sun and end at the blue dots on the right side of Earth.
• Only **one** line should be drawn from a particular point.
• Not all of the blue dots need to have lines between them.
moon
Delete
Add Point
Connect Line


Grid Interaction (GI Click up/Add Arrow) Example
Students investigate how the direction of light affects plant growth. They grow three plants in individual cardboard boxes using light from lamps. The picture shows the growth of Plant 1 with light coming from directly above the plant.
Plant 1
The students want to set up Plant 2 and Plant 3 with a light source to complete the investigation.
A. Click on one blank circle for Plant 2 and one blank circle for Plant 3 to show the direction of the light source for each plant to complete the investigation.
B. Use the Add Arrow button to draw an arrow showing the predicted growth of Plant 2 and Plant 3 based on the light source on each plant.
• Draw only one arrow for Plant 2.
• Draw only one arrow for Plant 3.
• There may be more than one correct answer.
Delete
Add Arrow
Experimental Setup
Plant 2
Plant 3


Simulation (SIM Nonscoring) Example
12
Students are studying different kinds of plants and the conditions that they grow in. They have planted four kinds of young plants.
Design and run an experiment that will show the effects of different amounts of sunlight and water on the plants.
Amount of Water
Little
Amount of Light
Direct Sun
Start
Agave
Moss
Rose
Fern
Amount of Water
Amount of Light
Agave
Moss
Rose
Fern
13
Which of the plants would grow best in a desert environment?
Agave
Fern
Moss
Rose
14
Which two kinds of plants could grow in the same environment based on the data from the experiment?
Agave and fern
Fern and moss
Moss and rose
Rose and agave
15
A student records some notes in a notebook during the experiment. Some of the notes are observations and some are inferences.
Select a box to identify whether each note is an observation or an inference.
Observation
Inference
Agave is a desert plant.
No type of fern can survive in direct sun.
The rose did not grow taller in the shade.
The fern turned brown when there was little water.


Simulation (SIM Scoring) Example
16
Students conducted a variety of experiments to understand how electricity flows to create light.
Design and run experiments to identify the effect of Mystery Component 4 on the other circuit components.
Circuit Component
Light Bulb
Mystery Component
Mystery Component 1
Start
Clear All Rows
Circuit Component
Mystery Component
Observations
Battery


Natural Language (NL) Example
The picture shows a manatee.
A. State one observation that can be made about the manatee from this picture. Be sure to identify it as an observation.
B. State one inference that can be made about the manatee from this picture. Be sure to identify it as an inference.
Type your answer in the space provided.

Selected Response (SR) Interactions

Selected Response interactions provide response options and the student selects the response(s).

SR Interaction Type	Task Demands that can be Assessed
Multiple Choice (MC)	Identify, Choose, Select, Label
Multi Select (MS)	Identify, Choose, Select, Label
Table Match (MI)	Classify, Categorize, Organize, Rank, Sort, Sequence
Editing Task Choice (ETC)	Classify, Categorize, Organize, Sort, Sequence, Compare, Label, Construct an explanation/argument, Describe, Summarize, Complete
Hot Text Selectable (HT)	Highlight, Identify, Select, Choose

Machine Scored Constructed Response (MSCR) Interactions

Machine Scored Constructed Response interactions require scoring logic or a machine rubric within the interaction. MSCR interaction types include:

Machine Scored Constructed Response Interaction Type	Task Demands that can be Assessed
Equation Editor (EQ)	Calculate, Mathematically describe/represent/model, Identify
Table Input (TI)	Calculate, Sequence, Identify, Organize, Chart
Grid Interaction (GI)	Graph, Model, Represent, Show, Create
Simulation Interaction (Sim)	Investigate, Experiment, Observe, Gather/collect data, Model
Natural Language (NL)	Describe, Compare, Summarize, Explain
Editing Task (ET)	Correct
Word Builder (WB)	Identify

Appendix 2-B

Item Specifications – Grade 3 – High School

Introduction

This document presents *cluster specifications* for use with the Next Generation Science Standards (NGSS). These standards are based on the Framework for K-12 Science Education. The present document is not intended to replace the standards, but rather to present guidelines for the development of items and item clusters used to measure those standards.

The remainder of this section provides a very brief introduction to the standards and the framework, an overview of the design and intent of the item clusters, and a description of the cluster specifications that follow. The bulk of the document is composed of cluster specifications, organized by grade and standard.

Background on the framework and standards

The Framework for K-12 Science Education are organized around three core dimensions of scientific understanding. The standards are derived from these same dimensions:

- Disciplinary Core Ideas: The fundamental ideas that are necessary for understanding a given science discipline. The core ideas all have broad importance within or across science or engineering disciplines, provide a key tool for understanding or investigating complex ideas and solving problems, relate to societal or personal concerns, and can be taught over multiple grade levels at progressive levels of depth and complexity.

- Science and Engineering Practices: The practices are what students DO to make sense of phenomena. They are both a set of skills and a set of knowledge to be internalized. The SEPs reflect the major practices that scientists and engineers use to investigate the world and design and build systems.

- Cross-Cutting Concepts: These are concepts that hold true across the natural and engineered world. Students can use them to make connections across seemingly disparate disciplines or situations, connect new learning to prior experiences, and more deeply engage with material across the other dimensions. The NGSS requires that students explicitly use their understanding of the CCCs to make sense of phenomena or solve problems.

- There is substantial overlap between and among the three dimensions. For example, the cross-cutting concepts are echoed in many of the disciplinary core ideas. The core ideas are often closely intertwined with the practices. This overlap reflects the nature of science itself. For example, we often come to understand and communicate causal relationships by employing models to make sense of observations. Even within a dimension, overlap exists. Quantifying characteristics of phenomena is important in developing an understanding of them, so employing computational and mathematical thinking in the construction and use of models is a very common scientific practice, and one of the cross-cutting concepts suggests that scientists often infer causality by observing patterns. In short, the dimensions are not orthogonal.

The framework envisions effective science education as occurring at the intersection of these interwoven dimensions: students learn science by doing science—applying the practices through the lens of the cross-cutting concepts to investigate phenomena that relate to the content of the disciplinary core ideas.

Item clusters

Each item cluster is designed to engage the examinee in a grade-appropriate, meaningful scientific activity aligned to a specific standard.

Each cluster begins with a phenomenon, an observable fact or design problem that engages student interest and can be explained, modeled, investigated, or designed using the knowledge and skill described by the standard in question. What it means to be observable varies across practices. For example, a phenomenon for a performance expectation exercising the a*nalyze data* practice may be observable through regularities in a data set, while standards related to the *development and use of models* might be something that can be watched, seen, felt, smelled, or heard.

What it means to be observable also varies across grade levels. For example, elementary-level phenomena are very concrete and directly observable. At the high school level, an observation of the natural world may be more abstract--for example, "observing" changes in the chemical composition of cells through the observation of macroscopic results of those changes on organism physiology, or through the measurement of system- or organ-level indications.
Content limits refine the intent of the performance expectations and provide limits on what may be asked of items in the cluster to structure the student activity. The content limits also reflect the disciplinary core ideas learning progressions that are present in the K-12 Framework for Science Education.

The task or goal should be explicitly stated in the stimulus or the first item in the cluster: statements such as "In the questions that follow, you will develop a model that will allow you to identify moons of Jupiter," or "In the questions below, you will complete a model to describe the processes that lead to the steam coming out of the teapot."
Whereas item clusters have been described elsewhere as "scaffolded," they are better described as providing structure to the task. For example, some clusters begin with students summarizing data to discover patterns that may have explanatory value. Depending on the grade level and nature of the standard, items may provide complete table shells or labeled graphs to be drawn, or may require the student to choose what to tabulate or graph. Subsequent items may ask the student to note patterns in the tabulated or graphed data and draw on domain content knowledge to posit explanations for the patterns.
These guidelines for clusters do not appear separately in the specifications. Rather, they apply to all clusters.

Structure of the cluster specifications

The item cluster specifications are designed to guide the work of item writers and the review of item clusters by stakeholders.
Each item cluster has the following elements:

- The text of the performance expectations, including the practice, core idea, and cross-cutting concept.
- Content limits, which refine the intent of the performance expectations and provide limits of what may be asked of examinees. For example, they may identify the specific formulae that students are expected to know or not know.
- Vocabulary, which identifies the relevant technical words that students are expected to know, and related words that they are explicitly not expected to know. The vocabulary included in both sections (Students are Expected to Know and Students Are Not Expected to Know) were developed after the reviews of standards at the current/preceding grades, the original NGSS documentation, and item writer reference documentation including the Children Writers' Word Book and ED Core Vocabularies in Reading, Mathematics, Science and Social Studies. All vocabulary included in the specifications was reviewed and edited by teacher committees during the specification reviews by states. Of course, the latter category should not be considered exhaustive, since the boundaries of relevance are ambiguous, and the list is limited by the imagination of the writers/reviewers.
- Sample phenomena, which provide some examples of the sort of phenomena that would support effective item clusters related to the standard in question. In general, these should be guideposts, and item writers should seek comparable phenomena, rather than drawing on those within the documents. Novelty is valued when applying scientific practices.
- Task demands comprise the heart of the specifications. These statements identify the types of items and activities that item writers should use, and each item written should be clearly linked to one or more of the demands. The verbs in the demands (e.g., *select, identify, illustrate, describe*) provide guidance on the types of interactions that item writers might employ to elicit the student response. We avoid explicitly identifying interaction types or item formats to accommodate future innovations and to avoid discouraging imaginative work by the item writers. All task demands should be aligned to a minimum of one of the three dimensions (DCI, SEP and CCCs) and across task demands within a cluster, all three dimensions must be addressed.

- For each cluster we present, the printed documentation includes the cluster, the task demands represented by each item, and its linkage to the content, practice and cross-cutting concept identified in the performance expectation.

Item cluster specifications follow, organized by domain and standard.

<table>
<tr><td>Performance Expectation</td><td colspan="3">3-PS2-1
Plan and conduct an investigation to provide evidence of the effects of balanced and unbalanced forces on the motion of an object.</td></tr>
<tr><td>Dimensions</td><td>Planning and Carrying Out Investigations
• Plan and conduct an investigation collaboratively to produce data to serve as the basis for evidence, using fair tests in which variables are controlled and the number of trials considered.</td><td>PS2.A: Forces and Motion
• Each force acts on one particular object and has both strength and a direction. An object at rest typically has multiple forces acting on it, but they add to give zero net force on the object. Forces that do not sum to zero can cause changes in the object’s speed or direction of motion. (Boundary: Qualitative and conceptual, but not quantitative addition of forces are used at this level.)

PS2.B: Types of Interactions
• Objects in contact exert forces on each other.</td><td>Cause and Effect
• Cause and effect relationships are routinely identified.</td></tr>
<tr><td>Clarifications and Content Limits</td><td colspan="3">Clarification Statements
• Examples could include an unbalanced force on one side of a ball can make it start moving, and balanced forces pushing on a box from both sides will not produce any motion at all.

Content Limits
• Assessment is limited to gravity being addressed as a force that pulls objects down.
• Assessment is limited to one variable at a time: number, size, or direction of forces.
• Assessment does include normal force, but not by name or magnitude.
• Assessment does not include quantitative force size, only qualitative and relative.</td></tr>
<tr><td>Science Vocabulary Students are Expected to Know</td><td colspan="3">Strength, direction, speed, gravity, net, sum, weight (physical).</td></tr>
<tr><td>Science Vocabulary Students are Not Expected to Know</td><td colspan="3">Velocity, acceleration, mass, friction, vector, quantitative, relative, scale, weight (mass • gravity), Newtons, normal force.</td></tr>
<tr><td colspan="4" align="center">Phenomena</td></tr>
<tr><td>Context/ Phenomena</td><td colspan="3">Example Phenomena for 3-PS2-1:
• Kids of the same size and strength play a game of tug of war. When the same number of kids are on each side, a ribbon tied to the rope does not move. When more kids are on one side, the rope moves in that direction.
• A ball rests on the ground, unmoving. When it is gently kicked, it moves slowly in the direction it was kicked. When it is kicked harder, it moves more quickly in the direction it was kicked.
• A box is sitting in the center of a table. Strings attached to the left and right sides of the box hang over the sides of the table. Identical weights can be attached to the end of these strings.
• A flat track with posts and rubber bands on either ends of the track. The student can pull a car back different distances to gather data.</td></tr>
</table>

This Performance Expectation and associated Evidence Statements support the following Task Demands.
Task Demands
1. Assemble, complete, or identify, from a collection including distractors, the essential components of an investigation that studies balanced and unbalanced forces on an object at rest and/or in motion.
2. Identify the variables in the investigation that are held constant and which are changing, and define important factors in the design including number of trials, methods, and techniques.
3. Identify the observations that should be collected in an investigation of an object's motion to determine the forces on the object and the causes of those forces.
4. Observe, collect, and record data from observations of the forces acting on an object at rest and/or in motion after forces of different strengths and/or directions are applied, including both balanced and unbalanced forces.*
5. Identify from a list, including distractors, the effects of forces on an object's motion and the cause of those forces.
6. Make predictions about the effects of changes in the motion of an object given specific forces. Predictions can be made by manipulating components of the investigation, completing illustrations, or selecting from lists with distractors.

*denotes those task demands which are deemed appropriate for use in stand-alone item development

<table>
<tr><td>Performance Expectation</td><td colspan="3">3-PS2-2
Make observations and/or measurements of an object’s motion to provide evidence that a pattern can be used to predict future motion.</td></tr>
<tr><td>Dimensions</td><td>Planning and Carrying Out Investigations
• Make observations and/or measurements to produce data to serve as the basis for evidence for an explanation of a phenomenon or to test a design solution.</td><td>PS2.A Forces and Motion
• The patterns of an object’s motion in various situations can be observed and measured; when that past motion exhibits a regular pattern, future motion can be predicted from it.</td><td>Patterns
• Patterns of change can be used to make predictions.</td></tr>
<tr><td>Clarifications and Content Limits</td><td colspan="3">Clarification Statements
• Examples of motion with a predictable pattern could include a child swinging in a swing, a ball rolling back and forth in a bowl, and two children on a seesaw.

Content Limits
• Technical terms, such as magnitude, velocity, momentum, and vector quantity, are not introduced at this level, but the concept that some quantities need both size and direction to be described is developed.
• <u>Students do not need to know:</u> Newton’s laws of motion, Law of Conservation of Energy</td></tr>
<tr><td>Science Vocabulary Students Are Expected to Know</td><td colspan="3">Speed, distance, height, time, mass, force, gravity, electrical field, static electricity, distribution of charged particles, electrical charge, negatively charged, positively charged, neutrally charged, magnetic field, polarity (magnetic), North pole, South pole, attraction, repulsion, electromagnet.</td></tr>
<tr><td>Science Vocabulary Students Are Not Expected to Know</td><td colspan="3">Frequency, amplitude, displacement, equilibrium position, oscillate, momentum, velocity, vector, elastic collision, inelastic collision, friction, acceleration of gravity, work, power, controlled variable, dependent variable, independent variable, kinetic energy, potential energy</td></tr>
<tr><td colspan="4">Phenomena</td></tr>
<tr><td>Context/ Phenomena</td><td colspan="3">Some example phenomena for 3-PS2-2:
• A boy and a girl play on a swing set. In 10 tries, the girl cannot get the boy to swing higher than the height she released him.
• A ball can be thrown farther when a person launches the ball from a plastic ball thrower rather than from his/her bare hand.
• A marble is rolled down a slide. It takes five seconds for the marble to reach the bottom of the slide. The same marble is rolled down another slide. This time, it takes the marble two seconds to reach the bottom of the slide.</td></tr>
<tr><td colspan="4">This Performance Expectation and associated Evidence Statements support the following Task Demands.</td></tr>
<tr><td colspan="4">Task Demands</td></tr>
<tr><td colspan="4">1. Identify the output data that should be collected in an investigation of an object’s motion.</td></tr>
<tr><td colspan="4">2. Make and/or record observations about an object’s motion as it repeats a pattern over time.</td></tr>
</table>

3. Generate or construct graphs, tables, assemblages of illustrations and/or labels of data that highlight patterns, trends, or correlations in the pattern of an object's motion. This may include sorting out distractors.*
4. Summarize data to highlight trends, patterns, or correlations in the motion of an object.
5. Use relationships identified in the data to predict/infer the future motion of an object.*
6. Identify patterns or evidence in the data that supports predictions/inferences about an object's future motion.*

*denotes those task demands which are deemed appropriate for use in stand-alone item development

<table>
<tr><td>Performance Expectation</td><td colspan="3">3-PS2-3
Ask questions to determine cause and effect relationships of electric or magnetic interactions between two objects not in contact with each other.</td></tr>
<tr><td>Dimensions</td><td>Asking Questions and Defining Problems
• Ask questions that can be investigated based on patterns such as cause and effect relationships.</td><td>PS2.B: Types of Interactions
• Electric and magnetic forces between a pair of objects do not require that the objects be in contact. The sizes of the forces in each situation depend on the properties of the objects and their distances apart and, for forces between two magnets, on their orientation relative to each other.</td><td>Cause and Effect
• Cause and effect relationships are routinely identified, tested, and used to explain change.</td></tr>
<tr><td>Clarifications and Content Limits</td><td colspan="3">Clarification Statements
• Examples of an electric force could include the force on hair from an electrically charged balloon and the electrical forces between a charged rod and pieces of paper; examples of a magnetic force could include the force between two permanent magnets, the force between an electromagnet and steel paperclips, and the force exerted by one magnet versus the force exerted by two magnets.
• Examples of cause and effect relationships could include how the distance between objects affects strength of the force and how the orientation of magnets affects the direction of the magnetic force.

Content Limits
• Assessment is limited to forces produced by objects that can be manipulated by students, and electrical interactions are limited to static electricity. Limit to strictly qualitative observations.
• Limit content to ask questions about how electric and magnetic objects interact, and the investigation of these phenomena within the scope of the classroom.
• Students should be able to identify the direction of the force, but not the shape of the magnetic or electric field.</td></tr>
<tr><td>Science Vocabulary Students are Expected to Know</td><td colspan="3">Attraction, repulsion, north pole, south pole, positive charge, negative charge, static electricity.</td></tr>
<tr><td>Science Vocabulary Students are Not Expected to Know</td><td colspan="3">Force fields, test charge, protons, neutrons, electrons, field gradients, insulator, conductor.</td></tr>
<tr><td colspan="4">Phenomena</td></tr>
<tr><td>Context/ Phenomena</td><td colspan="3">Example Phenomena for 3-PS2-3:
• A balloon rubbed against a sweater attracts a whole grain oat O-shaped cereal attached to a string.
• A magnet floats on top of another magnet when aligned correctly.
• A magnet touching the underside of a glass table can move a piece of metal sitting above it on top of the table.
• Two opposite poled magnets suspended by strings in air will levitate.</td></tr>
</table>

This Performance Expectation and associated Evidence Statements support the following Task Demands.
Task Demands
1. Select or identify from a collection, questions that will help clarify the properties that are correlated with the strength or direction of the forces in the phenomenon. In addition to plausible distractors, distractors may also include non-testable (“nonscientific”) questions.*
2. Make and/or record observations about how the size of the forces, both magnetic and electric, depend on different characteristics such as strength/orientation of the magnet, the amount of electric charge, materials, etc.
3. Identify, describe, or select from a collection, characteristics, properties, features, and/or processes to be manipulated or held constant, while gathering information to answer a well-articulated question about the cause and effect relationships of electric or magnetic interactions.*
4. Select or describe conclusions relevant to the question posed which are supported by the data, especially inferences about causes and effects, related to static electricity and/or magnetism.
5. Predict outcomes when properties or proximity of the objects are changed, given the inferred cause and effect relationships, related to static electricity and/or magnetism.

*denotes those task demands which are deemed appropriate for use in stand-alone item development
TD1 and TD3 **must be used together.

<table>
<tr><td>Performance Expectation</td><td colspan="3">3-PS2-4
Define a simple design problem that can be solved by applying scientific ideas about magnets.</td></tr>
<tr><td>Dimensions</td><td>Asking Questions and Defining Problems
• Define a simple problem that can be solved through the development of a new or improved object or tool.</td><td>PS2.B: Types of Interactions
• Electric and Magnetic forces between a pair of objects do not require the objects to be in contact. The sizes of the forces in each situation depend on the properties of the objects and their distances apart, and, for forces between two magnets, on their orientation relative to each other.</td><td></td></tr>
<tr><td>Clarifications and Content Limits</td><td colspan="3">Clarification Statements
• Examples of problems could include constructing a latch to keep a door shut and creating a device to keep two moving objects from touching each other.

Content Limits
• Students only need to know the basics about magnets. They do not need to know about the magnetic field and how it is shaped for different objects, etc.
• Students do not need to know how a magnet can magnetize other objects; they just need to know that it does. For example, a paper clip is not magnetic but will be attracted to a magnet. (The student does not need to know anything about magnetic domains.)
• Students do not need to know how electricity and magnetism are coupled (that moving electrons create a magnetic field and that a changing magnetic field creates a current).
• Students do not need to know anything about magnets except that they can repel/attract each other based on their orientation relative to each other.</td></tr>
<tr><td>Science Vocabulary Students Are Expected to Know</td><td colspan="3">Magnetic, attraction, repulsion, non-contact force, pole, North Pole, South Pole, bar magnet.</td></tr>
<tr><td>Science Vocabulary Students Are Not Expected to Know</td><td colspan="3">Force fields, field gradients, conductor, orientation, magnetic field, exert, interaction, electromagnetism.</td></tr>
<tr><td colspan="4">Phenomena</td></tr>
<tr><td>Context/ Phenomena</td><td colspan="3">Some example phenomena for 3-PS2-4:
• The shower leaks because the curtain is not secured to the bottom of the bathtub.
• Things continually fall out of a handbag because the latch is not secure.
• While working on a project, pencil shavings were dropped on the carpet and the vacuum may not have cleaned them all up.
• Two carts used in experiments keep damaging each other when they collide.</td></tr>
<tr><td colspan="4">This Performance Expectation and associated Evidence Statements support the following Task Demands.</td></tr>
<tr><td colspan="4">Task Demands</td></tr>
<tr><td colspan="4">1. Identify or assemble from a collection, including distractors, the relevant aspects of the problem that given design solutions, if implemented, will resolve/improve.</td></tr>
</table>

2. Articulate, describe, illustrate, or select the relationships, interactions, and/or processes to be explained OR to be used to solve the problem. This may entail sorting relevant from irrelevant information or features.
3. Express or complete a causal chain explaining how the repulsion or attraction of magnets will solve the problem that has been identified. This may include indicating directions of causality in an incomplete model such as a flow chart or diagram, or completing cause-and-effect chains.
4. Using given data, propose/illustrate/assemble a potential device (prototype) or solution.
5. Describe, identify, and/or select information needed to support an explanation about the proposed solution.

<table>
<tr><td>Performance Expectation</td><td colspan="3">3-LS1-1
Develop models to describe that organisms have unique and diverse life cycles but all have in common birth, growth, reproduction, and death.</td></tr>
<tr><td>Dimensions</td><td>Developing and Using Models
• Develop models to describe phenomenon.</td><td>LS1.B: Growth and Development of Organisms
• Reproduction is essential to the continued existence of every kind of organism. Plants and animals have unique and diverse life cycles.</td><td>Patterns
• Patterns of change can be used to make predictions.</td></tr>
<tr><td>Clarifications and Content Limits</td><td colspan="3">Clarification Statements
• Changes organisms go through during their lifetime form a pattern.

Content Limits
• Assessment of plant life cycles is limited to those of flowering plants.
• Assessment does not include details of human reproduction.
• <u>Students do not need to know:</u> the alternation of generations life cycle, the human reproductive system, mitosis and meiosis.</td></tr>
<tr><td>Science Vocabulary Students Are Expected to Know</td><td colspan="3">Adult, growth, parent, pollen, offspring, structure, feature, trait, young, root, stem, leaf/leaves, seed, flower, petal</td></tr>
<tr><td>Science Vocabulary Students Are Not Expected to Know</td><td colspan="3">Organism, breed, transfer, development, germination, reproductive system, cell, tissue, egg, fertilize, genetic, unicellular, multicellular, specialized cell, cell division, variation, juvenile, metamorphosis, chrysalis, pupa, spores, pistil, stamen, ovary, anther, filament, sepal, receptacle, ovule, stigma, style.</td></tr>
<tr><td colspan="4">Phenomena</td></tr>
<tr><td>Context/ Phenomena</td><td colspan="3">Some example phenomena for 3-LS1-1:
• A young moth builds a soft case around it called a cocoon and a young butterfly builds a hard case called a chrysalis.
• A young ladybug looks very different from an adult ladybug.
• Plants and animals both form eggs.
• A pea planted in the ground grows into a new pea plant.</td></tr>
<tr><td colspan="4">This Performance Expectation and associated Evidence Statements support the following Task Demands.</td></tr>
<tr><td colspan="4">Task Demands</td></tr>
<tr><td colspan="4">1. Select the components needed to model the phenomenon. Components might include stages of life cycles such as birth, growth, reproduction, and death.</td></tr>
<tr><td colspan="4">2. Assemble or complete an illustration or flow chart that is capable of representing the patterns in life cycles of different types of organisms.</td></tr>
<tr><td colspan="4">3. Manipulate the components of a model to demonstrate the changes, properties, processes and/or events that act to result in a phenomenon.</td></tr>
<tr><td colspan="4">4. Make predictions about the effects of changes in life cycles on organisms. Predictions can be made by manipulating model components, completing illustrations, or selecting from a list with distractors.</td></tr>
</table>

5. Given models or diagrams of life cycles, identify relevant components such as birth, growth, reproduction, and death, and how the life cycles are different in each scenario.
6. Identify missing components, relationships, or other limitations of the model of a life cycle.
7. Describe, select, or identify the relationships among components of a model that describe the patterns of life cycles among different organisms.

<table>
<tr><td>Performance Expectation</td><td colspan="3">3-LS2-1
Construct an argument that some animals form groups that help members survive.</td></tr>
<tr><td>Dimensions</td><td>Engaging in Argument from Evidence
• Construct an argument with evidence, data, and/or a model.</td><td>LS2.D: Social Interactions and Group Behavior
• Being part of a group helps animals obtain food, defend themselves, and cope with changes. Groups may serve different functions and vary dramatically in size.</td><td>Cause and Effect
• Cause and effect relationships are routinely identified and used to explain change.</td></tr>
<tr><td>Clarifications and Content Limits</td><td colspan="3">Clarification Statements
• Focus is on how being part of a group helps animals obtain food, defend themselves, and cope with changes, and does not cover how group behavior evolved as a result of a survival advantage.

Content Limits
• Assessment does not include the evolution of group behavior.
• <u>Students do not need to know</u>: social hierarchy in animal groups (pecking order, dominance, submissive, altruism).</td></tr>
<tr><td>Science Vocabulary Students Are Expected to Know</td><td colspan="3">Environment, prey, predator, characteristic, habitat, species, herd, inherit, trait, diet, mate, parent</td></tr>
<tr><td>Science Vocabulary Students Are Not Expected to Know</td><td colspan="3">Organism, social, relative, predation, hereditary, harmful, beneficial, variation, probability, adaptation, decrease, increase, behavioral, variation, ecosystem, pecking order, dominance/submissive behavior, hierarchy, migrate, defend.</td></tr>
<tr><td colspan="4">Phenomena</td></tr>
<tr><td>Context/ Phenomena</td><td colspan="3">Some example phenomena for 3-LS2-1:
• In Yellowstone National Park, a wolf preys on a much larger bison.
• In the Willamette Valley, a colony of beavers builds a dam.
• A colony of ants protects its nests.
• A male honey bee returns to a hive each day.</td></tr>
<tr><td colspan="4">This Performance Expectation and associated Evidence Statements support the following Task Demands.</td></tr>
<tr><td colspan="4">Task Demands</td></tr>
<tr><td colspan="4">1. Identify patterns or evidence in the data that support inferences and/or determine relationships about the effect of group membership on survival of an animal.</td></tr>
<tr><td colspan="4">2. Understand and generate simple bar graphs or tables that document patterns, trends, or relationships between group membership and survival.</td></tr>
<tr><td colspan="4">3. Sort observations/evidence into those that appear to support or not support an argument.</td></tr>
<tr><td colspan="4">4. Based on the provided data, identify or describe a claim regarding the relationship between survival of an animal and being a member of a group.</td></tr>
<tr><td colspan="4">5. Identify, summarize, select or organize given data or other information to support or refute a claim regarding the relationship between group membership and survival of an animal.*</td></tr>
</table>

6. Using evidence, explain the relationship between group membership and survival.*

*denotes those task demands which are deemed appropriate for use in stand-alone item development

This page was intentionally left blank.

<table>
<tr><td>Performance Expectation</td><td colspan="3">3-LS3-1
Analyze and interpret data to provide evidence that plants and animals have traits inherited from parents and that variation of these traits exists in a group of similar organisms.</td></tr>
<tr><td>Dimensions</td><td>Analyzing and Interpreting Data
• Analyze and interpret data to make sense of phenomena using logical reasoning.</td><td>LS3.A: Inheritance of Traits
• Many characteristics of organisms are inherited from their parents.

LS3.B: Variation of Traits
• Different organisms vary in how they look and function because they have different inherited information</td><td>Patterns
• Similarities and differences in patterns can be used to sort and classify natural phenomena.</td></tr>
<tr><td>Clarifications and Content Limits</td><td colspan="3">Clarification Statement
• Patterns are the similarities and differences in traits shared between offspring and their parents, or among siblings.

Content Limits
• Emphasis is on organisms other than humans.
• Assessment does not include genetic mechanisms of inheritance and prediction of traits, including concepts of dominant/recessive traits or sex-linked traits.
• Assessment is limited to non-human examples.
• Graphs and charts can include bar graphs, pictographs, pie charts, tally chart.
• Types of math can include simple fractions, simple addition/subtraction.</td></tr>
<tr><td>Science Vocabulary Students are Expected to Know</td><td colspan="3">Parent, sibling, characteristic, offspring, parent-offspring similarity, feature, inherit, inherited characteristic, reproduce</td></tr>
<tr><td>Science Vocabulary Students are Not Expected to Know</td><td colspan="3">Transfer, variation, allele, hereditary information, identical, Punnett square, transmission, gene, genetic, genetic variation, dominant trait, recessive trait.</td></tr>
<tr><td colspan="4">Phenomena</td></tr>
<tr><td>Context/ Phenomena</td><td colspan="3">For this performance expectation the phenomena are sets of data. Those are the observed facts that students will look at to discover patterns. Below, we enumerate some of the patterns that might comprise the data sets (phenomena) to be analyzed.

Example Phenomena for 3-LS3-1:
• Two corn plants in a garden reproduce. In the next generation, the offspring vary in height. (Augmentation: We will provide a data table displaying each member of the subsequent generation and the relevant trait possessed.)
• Over a four-year period, the offspring of two tall blueberry plants always grow taller than the offspring of two nearby short blueberry plants. (Augmentation: We will provide a data table of the number of offspring of each plant height over a four-year period, correlated with the parent plants.)</td></tr>
<tr><td colspan="4">This Performance Expectation and associated Evidence Statements support the following Task Demands.</td></tr>
<tr><td colspan="4">Task Demands</td></tr>
</table>

1. Organize or summarize data to highlight trends, patterns, or correlations between the traits of offspring and those of their parents and/or siblings.*
2. Generate graphs or tables that document patterns, trends, or correlations in inheritance of traits.*
3. Identify patterns or evidence in the data that support inferences about inheritance of traits from parents to offspring.*

*denotes those task demands which are deemed appropriate for use in stand-alone item development

<table>
<tr><td>Performance Expectation</td><td colspan="3">3-LS3-2
Use evidence to support the explanation that traits can be influenced by the environment.</td></tr>
<tr><td>Dimensions</td><td>Constructing explanations and designing solutions
• Use evidence (e.g., observations, patterns) to support an explanation.</td><td>LS3.A: Inheritance of Traits
• Other characteristics result from individuals' interactions with the environment, which can range from diet to learning. Many characteristics involve both inheritance and environment.

LS3.B: Variation of Traits
• The environment also affects the traits that an organism develops.</td><td>Cause and Effect
• Cause-and-effect relationships are routinely identified and used to explain change.</td></tr>
<tr><td>Clarifications and Content Limits</td><td colspan="3">Clarification Statements
• Examples of the environment affecting a trait could include normally tall plants that are grown with insufficient water and are stunted; and, a pet dog that is given too much food and little exercise and becomes overweight.

Content Limits
• Assessment should focus on physical traits.
• Content should not include human traits.</td></tr>
<tr><td>Science Vocabulary Students Are Expected to Know</td><td colspan="3">Offspring, feature, inherit, diet, survival, flood, drought, habitat, reproduce</td></tr>
<tr><td>Science Vocabulary Students Are Not Expected to Know</td><td colspan="3">Organism, variation, version, harmful, beneficial, increase, decrease, trend</td></tr>
<tr><td colspan="4">Phenomena</td></tr>
<tr><td>Context/ Phenomena</td><td colspan="3">Some example phenomena for 3-LS3-2:
• The arctic fox is white in winter but turns brown in the summer.
• Flamingoes are born gray, but some become very pink as they grow.
• Trees growing on the edge of cliffs are often bent.
• A goldfish in a pond grows larger than one in a fish bowl.</td></tr>
<tr><td colspan="4">This Performance Expectation and associated Evidence Statements support the following Task Demands.</td></tr>
<tr><td colspan="4">Task Demands</td></tr>
<tr><td colspan="4">1. Describe or select the relationships, interactions, or processes to be explained. This may entail sorting relevant from irrelevant information or features.</td></tr>
<tr><td colspan="4">2. Express or complete a causal chain explaining that traits can be influenced by the environment. This may include indicating directions of causality in an incomplete model such as a flow chart or diagram, or completing cause-and-effect chains.</td></tr>
<tr><td colspan="4">3. Identify evidence supporting the inference of causation that is expressed in a causal chain.</td></tr>
</table>

4. Use an explanation to predict changes in the trait of an organism given a change in environmental factors.
5. Describe, identify, and/or select information needed to support an explanation of environmental influence on traits.

This page was intentionally left blank.

<table>
<tr><td>Performance Expectation</td><td colspan="3">3-LS4-1
Analyze and interpret data from fossils to provide evidence of the organisms and the environments in which they lived long ago.</td></tr>
<tr><td>Dimensions</td><td>Analyzing and Interpreting Data
• Analyze and interpret data to make sense of phenomena using logical reasoning, mathematics, and/or computation.</td><td>LS4.A: Evidence of Common Ancestry and Diversity
• Some kinds of plants and animals that once lived on Earth are no longer found anywhere.
• Fossils provide evidence about the types of organisms that lived long ago and also about the nature of their environments.</td><td>Scale, Proportion, and Quantity
• Observable phenomena exist from very short to very long periods.</td></tr>
<tr><td>Clarifications and Content Limits</td><td colspan="3">Clarification Statements
• Examples of data could include type, size, and distributions of fossil organisms.
• Examples of fossils and environments could include marine fossils found on dry land, tropical plant fossils found in Arctic areas, and fossils of extinct organisms.
• Focus is on the fossils and environment in which the organisms lived, not how the fossils got to where they are today.
• Data can be represented in tables and/or various graphic displays.
• Data collected by different groups can be compared and contrasted to discuss similarities and differences in their findings.

Content Limits
• Assessment does not include identification of specific fossils or present plants and animals.
• Assessment is limited to major fossil types and relative ages.
• Graphs and charts can include bar graphs, pictographs, pie charts, and tally charts.
• Types of math can include simple addition/subtraction.
• Standard units that can be used to measure and describe physical quantities such as weight, time, temperature, and volume.</td></tr>
<tr><td>Science Vocabulary Students Are Expected to Know</td><td colspan="3">Exist, existence, ecosystem, characteristic, habitat, species, volcanic eruption, climate, extinct, extinction, predator, time period, earthquake, erosion, weathering.</td></tr>
<tr><td>Science Vocabulary Students Are Not Expected to Know</td><td colspan="3">Chronological order, fossil record, radioactive dating, descent, ancestry, evolution, evolutionary, genetic, relative, rock layer.</td></tr>
<tr><td colspan="4">Phenomena</td></tr>
<tr><td>Context/ Phenomena</td><td colspan="3">For this performance expectation, the phenomena are sets of data. Those are the observed facts that students will look at to discover patterns. Below, we enumerate some of the patterns that might comprise the data sets (phenomena) to be analyzed.

Some example phenomena for 3-LS4-1:
• Fossil trees are found in sedimentary rocks in Antarctica.
• The Redwall Limestone in the Grand Canyon contains many different fossils including corals, clams, octopi, and fish.</td></tr>
</table>

	• Whale fossils have been found in rocks in the Andes Mountains. • Fossils of corals and snails are found in Iowa.
This Performance Expectation and associated Evidence Statements support the following Task Demands.	
Task Demands	
1. Organize or summarize data to highlight trends, patterns, or correlations between plant and animal fossils and the environments in which they lived.	
2. Generate graphs or tables that document patterns, trends, or correlations in the fossil record.	
3. Identify evidence in the data that supports inferences about plant and animal fossils and the environments in which they lived.	

<table>
<tr><td>Performance Expectation</td><td colspan="3">3-LS4-2
Use evidence to construct an explanation for how the variations in characteristics among individuals of the same species may provide advantages in surviving, finding mates, and reproducing.</td></tr>
<tr><td>Dimensions</td><td>Constructing Explanations and Designing Solutions
• Use evidence (e.g., observations, patterns) to construct an explanation.</td><td>LS4.B: Natural Selection
• Sometimes the differences in characteristics between individuals of the same species provide advantages in surviving, finding mates, and reproducing.</td><td>Cause and Effect
• Cause and effect relationships are routinely identified and used to explain change.</td></tr>
<tr><td>Clarifications and Content Limits</td><td colspan="3">Clarification Statements
• Examples of cause and effect relationships could be: plants that have larger thorns than other plants may be less likely to be eaten, and animals that have better camouflage coloration than other animals may be more likely to survive and therefore more likely to leave offspring.

Content Limits
• Differences between individuals helping or hurting chances of survival and reproduction should be included.
• Data sets can include not only common trends but also outliers and anomalous data points.
• Analysis of data should be limited to patterns and trends.
• Students are not expected to evaluate the extent at which the sample is representative of a population.
• <u>Students do not need to know</u>: Mechanisms or patterns of inheritance, detailed life cycles.</td></tr>
<tr><td>Science Vocabulary Students are Expected to Know</td><td colspan="3">Variation, advantage, reproduce, relationship, mating, breeding, behavior, plumage, pollination, camouflage.</td></tr>
<tr><td>Science Vocabulary Students are Not Expected to Know</td><td colspan="3">Natural and artificial selection, evolution, genetics, adaptation,</td></tr>
<tr><td colspan="4">Phenomena</td></tr>
<tr><td>Context/ Phenomena</td><td colspan="3">Some example phenomena for 3-LS4-2:
• The same species of walking stick in California has two different color variations. The green walking sticks are found on bushes with thick green leaves, whereas the striped walking sticks are found on bushes with needle-like leaves.
• In a given population, there are more male [X Bird] with larger, brighter feathers in the population than males with smaller, muted feathers.
• Acacia trees that are browsed upon by X animal grow longer thorns at X height. Acacia trees that are browsed upon by Y animal grow longer thorns at Y height. Acacia trees that are not browsed upon at all do not grow longer thorns.
• Io moths use eyespots on their inner wings to frighten predators away. Larger eyespots are more effective.</td></tr>
<tr><td colspan="4">This Performance Expectation and associated Evidence Statements support the following Task Demands.</td></tr>
</table>

Task Demands
1. Articulate, describe, illustrate, or select the variations of characteristics to be explained. This may entail sorting relevant from irrelevant information or features.
2. Identify evidence supporting the conclusion that the variations in characteristics among individuals of the same species may provide advantages in surviving, finding mates, and reproducing.
3. Describe, identify, and/or select information needed to support an explanation that a characteristic provides advantages in surviving and reproducing.
4. Select or identify a prediction about survival or reproduction rates given a change in a characteristic. The prediction should follow from an explanation or causal relationship supported in earlier items.
5. Identify additional evidence that would help clarify, support, or contradict a hypothesized relationship between characteristics of individuals and their chances of survival and reproductive rates.
6. Express or complete a causal chain that explains how different characteristics among individuals of the same species provide advantages in survival and reproduction. This may include indicating directions of causality in an incomplete model such as a flow chart or diagram or completing cause and effect chains.*
7. Use evidence to construct an explanation for differences in survival and/or reproduction given a difference in traits between individuals of the same species.*

*denotes those task demands which are deemed appropriate for use in stand-alone item development

<table>
<tr><td>Performance Expectation</td><td colspan="3">3-LS4-3
Construct an argument with evidence that in a particular habitat some organisms can survive well, some survive less well, and some cannot survive at all.</td></tr>
<tr><td>Dimensions</td><td>Engaging in Argument from Evidence
• Construct an argument with evidence.</td><td>LS4.C: Adaptation
• For any particular environment, some kinds of organisms survive well, some survive less well, and some cannot survive at all.</td><td>Cause and Effect
• Cause and effect relationships are routinely identified and used to explain change.</td></tr>
<tr><td>Clarifications and Content Limits</td><td colspan="3">Clarification Statements
• Examples of evidence could include needs and characteristics of the organisms and habitats involved. The organisms and their habitat make up a system in which the parts depend on each other.

Content Limits
• While students are not expected to know the definitions to vocabulary terms such as extinction, climate, and mimic, they are expected to know the general concepts behind these terms.
• <u>Students do not need to know:</u> mechanisms of natural selection and evolution of species.</td></tr>
<tr><td>Science Vocabulary Students Are Expected to Know</td><td colspan="3">Habitat, health, species, population, region, resource, behavior, growth, petal, thorn, structure, characteristics, mate, trait.</td></tr>
<tr><td>Science Vocabulary Students Are Not Expected to Know</td><td colspan="3">Organism, threaten, impact, terrestrial, climate change, response, body plan, external, function, internal, invertebrate, adaptation, beneficial change, detrimental change, species diversity, gene, variation, artificial selection, natural selection.</td></tr>
<tr><td colspan="4">Phenomena</td></tr>
<tr><td>Context/ Phenomena</td><td colspan="3">Some example phenomena for 3-LS4-3:
• Desert plants are able to survive where there is little to no rain.
• Black bears survive the harsh winter months of their forest habitats by going into a deep sleep.
• The artic fox is better able to survive in colder climates than the red fox.
• Emperor penguins have special traits which help them survive in Antarctica.</td></tr>
<tr><td colspan="4">This Performance Expectation and associated Evidence Statements support the following Task Demands.</td></tr>
<tr><td colspan="4">Task Demands</td></tr>
<tr><td colspan="4">1. Organize or summarize data to highlight trends, patterns, and/or determine relationships between the traits of an organism and survival in its environment.</td></tr>
<tr><td colspan="4">2. Understand and generate simple bar graphs or tables that document patterns, trends, or relationships between traits of an organism and its survival in a particular environment.</td></tr>
<tr><td colspan="4">3. Identify patterns or evidence in the data that supports inferences about characteristics of an organism and those of its environment.</td></tr>
</table>

4. Based on the provided data, identify or describe a claim regarding the relationship between the characteristics of an organism and survival in a particular environment.*
5. Evaluate the evidence to sort relevant from irrelevant information regarding survival of an organism in a particular environment.

*denotes those task demands which are deemed appropriate for use in stand-alone item development

<table>
<tr><td>Performance Expectation</td><td colspan="3">3-LS4-4
Make a claim about the merit of a solution to a problem caused when the environment changes and the types of plants and animals that live there may change.</td></tr>
<tr><td>Dimensions</td><td>Engaging in Argument from Evidence
• Make a claim about the merit of a solution to a problem by citing relevant evidence about how it meets the criteria and constraints of the problem.</td><td>LS2.C: Ecosystem Dynamics, Functioning, and Resilience
• When the environment changes in ways that affect a place’s physical characteristics, temperature, or availability of resources, some organisms survive and reproduce, others move to new locations, yet others move into the transformed environment, and some die. (secondary)

LS4.D: Biodiversity and Humans
• Populations live in a variety of habitats and change in those habitats affects the organisms living there.</td><td>Systems and System Models
• A system can be described in terms of its components and their interactions.</td></tr>
<tr><td>Clarifications and Content Limits</td><td colspan="3">Clarification Statements
• Examples of environmental changes could include changes in land characteristics, water distribution, temperature, food, and other organisms.

Content Limits
• Assessment is limited to a single environmental change.
• Assessment does not include the greenhouse effect or climate change.
• <u>Students do not need to know</u>: greenhouse effect, ultraviolet (UV) radiation, nuclear disasters.</td></tr>
<tr><td>Science Vocabulary Students Are Expected to Know</td><td colspan="3">Population, organism, community, habitat, resource, reproduce, shelter, temperature, matter, predator, prey, flood, frost, tide</td></tr>
<tr><td>Science Vocabulary Students Are Not Expected to Know</td><td colspan="3">Ecosystem, biotic, abiotic, food web, producer, consumer, decomposer, photosynthesis, pollinate, adapt, energy flow, biosphere, sustain, predation, mutualism, carrying capacity, volcano, earthquake, drought, arid, blight.</td></tr>
<tr><td colspan="4">Phenomena</td></tr>
<tr><td>Context/ Phenomena</td><td colspan="3">Some example phenomena for 3-LS4-4
• To help ornamental bushes grow, no other plants should grow in their immediate vicinity.
• Before stocking a lake with fish, the lake pollution needs to be reduced.
• A late frost threatens the orange groves in Georgia.</td></tr>
<tr><td colspan="4">This Performance Expectation and associated Evidence Statements support the following Task Demands.</td></tr>
<tr><td colspan="4">Task Demands</td></tr>
<tr><td colspan="4">1. Articulate, describe, illustrate, or select the relationships, interactions, and/or processes involved when the types of plants and/or animals change as a result of environmental changes. This may entail sorting relevant from irrelevant information or features.</td></tr>
<tr><td colspan="4">2. Identify a problem that results when the types of plants and/or animals change as a result of environmental changes.</td></tr>
</table>

3. Express or complete a causal chain explaining a solution to problem that results when the types of plants and/or animals change as a result of environmental changes. The causal chain should include the ecosystem before the environmental change, the environmental change, the problem to plants and animals resulting from the environmental change, the solution to the problem, and the effect(s) of the solution on the ecosystem. This may include indicating directions of causality in an incomplete model such as a flow chart or diagram, or completing cause-and-effect chains.*
4. Identify and/or evaluate evidence related to a solution to a problem caused when the types of plants and/or animals change as a result of environmental changes. The evidence may support or refute the solution, or students may identify missing evidence.
5. Evaluate a solution to a problem that results when the types of plants and/or animals change as a result of environmental changes, including how the solution may affect plants, animals, and/or other aspects of the ecosystem.*
6. Identify information or data needed to support or refute a claim regarding a problem resulting from an environmental change affecting the native plants and animals.

*denotes those task demands which are deemed appropriate for use in stand-alone item development

<table>
<tr><td>Performance Expectation</td><td colspan="3">3-ESS2-1
Represent data in tables and graphical displays to describe typical weather conditions expected during a particular season.</td></tr>
<tr><td>Dimensions</td><td>Analyzing and Interpreting Data
• Represent data in tables and various graphical displays (bar graphs and pictographs) to reveal patterns that indicate relationships.</td><td>ESS2.D: Weather and Climate
• Scientists record patterns of the weather across different times and areas so that they can make predictions about what kind of weather might happen next.</td><td>Patterns
• Patterns of change can be used to make predictions.</td></tr>
<tr><td>Clarifications and Content Limits</td><td colspan="3">Clarification Statements
• Examples of data could include average temperature, precipitation, and wind direction.

Content Limits
• Assessment of graphical displays is limited to pictographs and bar graphs.
• Assessment does not include climate change.
• <u>Students do not need to know:</u> probabilities or how to calculate them, fronts and pressure systems, the movements of weather systems.</td></tr>
<tr><td>Science Vocabulary Students Are Expected to Know</td><td colspan="3">Season, weather, temperature, precipitation, patterns, average, latitude, longitude</td></tr>
<tr><td>Science Vocabulary Students Are Not Expected to Know</td><td colspan="3">Probability, anthropogenic change</td></tr>
<tr><td colspan="4">Phenomena</td></tr>
<tr><td>Context/ Phenomena</td><td colspan="3">Some example phenomena for 3-ESS2-1:
• Vienna, Austria, records more sunny days in the summer than in the winter. Data: Average sunshine hours by month for the city, given as a table or graph.
• People in Florida can often go outside without jackets during the winter. Data: Months and Temperatures for Florida, given as table or graph.
• Visitors to the desert in Death Valley, California, were surprised to be rained on. Data: Months and Precipitation Averages for the region given as table or graph.
• Flags in California’s San Joaquin Valley are seen blowing to the SE for most of the year, but are seen blowing to the NW in winter months. Data: Monthly average wind direction (and maybe speed) for the region, given as a table or graphic with wind direction arrows.</td></tr>
<tr><td colspan="4">This Performance Expectation and associated Evidence Statements support the following Task Demands.</td></tr>
<tr><td colspan="4">Task Demands</td></tr>
<tr><td colspan="4">1. Organize and/or arrange (e.g., using illustrations and/or labels), or summarize data to highlight trends, patterns, or correlations in weather patterns.*</td></tr>
<tr><td colspan="4">2. Generate/construct graphs, tables, or assemblages of illustrations and/or labels of data that document patterns, trends, or correlations in weather patterns. This may include sorting out distractors.*</td></tr>
</table>

3.	Use relationships and patterns identified in the data to predict weather.
4.	Identify patterns or evidence in the data that support conclusions about weather. ****

*denotes those task demands which are deemed appropriate for use in stand-alone item development.
**TD4 can be used for stand-alone item development if paired with TD2.

Performance Expectation	**3-ESS2-2** Obtain and combine information to describe climates in different regions of the world.		
Dimensions	**Obtaining, Evaluating, and Communicating Information** • Obtain and combine information from books and other reliable media to explain phenomena.	**ESS2.D: Weather and Climate** • Climate describes a range of an area's typical weather conditions and the extent to which those conditions vary over years.	**Patterns** • Patterns of change can be used to make predictions.
Clarifications and Content Limits	**Content Limits** • Students do not need to know: complex interactions that cause weather patterns and climate, the role of the water cycle in weather.		
Science Vocabulary Students Are Expected to Know	Prediction, precipitation, glacier, ocean, region, climate, vegetation, latitude, longitude, drought, temperature, freeze, atmosphere.		
Science Vocabulary Students Are Not Expected to Know	Average, high pressure, low pressure, air mass, altitude, humidity, radiation, water cycle.		
Phenomena			
Context/ Phenomena	Some example phenomena for 3-ESS2-2: • Anchorage, Alaska has cool summers and very cold winters with a lot of snowfall. • It often snows in Colorado in July, but it does not often snow in Kansas in July. • On the western side of the Cascade Mountains of Oregon, it rains frequently, but on the eastern side, it does not. • The temperature in London, England does not get very hot in summer or very cold in winter.		
This Performance Expectation and associated Evidence Statements support the following Task Demands.			
Task Demands			
1. Organize and/or arrange data (including labels and symbols) regarding the climates in different regions to highlight/identify trends or patterns, or make comparisons/contrasts between different regions and/or climatically relevant aspects of their geology and/or geography.*			
2. Generate or construct tables or assemblages of data (including labels and symbols) that document the similarities and differences between climates of different regions (this includes completing incomplete maps).			
3. Analyze and interpret scientific evidence (including textural and numerical information as well labels and symbols) from multiple sources (e.g., texts, maps, and/or graphs) that help identify patterns in weather in regions of different climate. This includes communicating the analysis or interpretation.*			
4. Analyze and interpret patterns of information on maps (including textural and numerical information as well labels and symbols) to explain, infer, or predict patterns of weather over time in a region.*			
5. Based on the information that is obtained and/or combined, identify, assert, describe, or illustrate a claim regarding the relationship between the location of a region and its climate, or the relationship between geological and/or geographical aspects/characteristics of a region and its climate.*			
6. Use spatial and/or temporal relationships identified in the obtained and/or combined climate data to predict typical weather conditions in a region.			

7. Organize and/or arrange data regarding the climate of a region to highlight/identify trends or relationships between the weather patterns of a region and its geology and/or geography.
8. Analyze and interpret scientific evidence (including textural and numerical information as well labels and symbols) from multiple sources (e.g., texts, maps, and/or graphs) that helps identify patterns in climate based on geography and/or geology. This includes communicating the analysis or interpretation.

*denotes those task demands which are deemed appropriate for use in stand-alone item development

<table>
<tr><td>Performance Expectation</td><td colspan="3">3-ESS3-1
Make a claim about the merit of a design solution that reduces the impacts of a weather-related hazard.</td></tr>
<tr><td>Dimensions</td><td>Engaging in Argument from Evidence
• Make a claim about the merit of a solution to a problem by citing relevant evidence about how it meets the criteria and constraints of the problem.</td><td>ESS3.B: Natural Hazards
• A variety of natural hazards result from natural processes. Humans cannot eliminate natural hazards but can take steps to reduce their impacts.</td><td>Cause and Effect
• Cause-and-effect relationships are routinely identified, tested, and used to explain change.</td></tr>
<tr><td>Clarifications and Content Limits</td><td colspan="3">Assessment Clarifications
• Examples of design solutions to weather-related hazards could include barriers to prevent flooding, wind resistant roofs, and lighting rods.</td></tr>
<tr><td>Science Vocabulary Students Are Expected to Know</td><td colspan="3">Natural process, earthquake, tsunami, tornado, flooding, severe weather, coastal erosion, landslide, avalanche, dams, levees, lightning, lightning rod, forecast, drought.</td></tr>
<tr><td>Science Vocabulary Students Are Not Expected to Know</td><td colspan="3">Fault line, names of clouds, names of storms, magma, types of volcanoes, low pressure, high pressure systems, El Niño, La Niña, jet stream.</td></tr>
<tr><td colspan="4">Phenomena</td></tr>
<tr><td>Context/ Phenomena</td><td colspan="3">For this performance expectation, phenomena should refer to hazard and one or more design solutions.

Some example phenomena for 3-ESS3-1:
• A building with a lightning rod is struck by lightning more often than the surrounding buildings.
• When the water level of the Feather River was high in February 2017, the water never rose higher than the levees around it, and no flooding occurred. When the water level of the Russian River was high in February 2017, the surrounding area flooded.
• A house built near the ocean in Surfside, New Jersey, sits on stilts/posts.
• A basement in a building fitted with a sump pump does not have mold while the basements of other nearby buildings have mold.</td></tr>
<tr><td colspan="4">This Performance Expectation and associated Evidence Statements support the following Task Demands.</td></tr>
<tr><td colspan="4">Task Demands</td></tr>
<tr><td colspan="4">1. Identify or assemble from a collection, including distractors, the relevant aspects of the hazard that a given design solution resolves/improves.</td></tr>
<tr><td colspan="4">2. Using the given information, select or identify the criteria against which the design solution should be judged.</td></tr>
<tr><td colspan="4">3. Using given information, select or identify constraints that the design solution must meet.</td></tr>
<tr><td colspan="4">4. Identify missing components, relationships, or other limitations of the design solution.</td></tr>
</table>

5. Use an explanation to predict the outcome of a hazard given a change in the design solution.
6. Make a claim about the merit of the design solution that can be defended.

<table>
<tr><td>Performance Expectation</td><td colspan="3">4-PS3-1
Use evidence to construct an explanation relating the speed of an object to the energy of that object.</td></tr>
<tr><td>Dimensions</td><td>Constructing Explanations and Designing Solutions
• Use evidence (e.g., measurements, observations, patterns) to construct an explanation.</td><td>PS3.A: Definitions of Energy
• The faster a given object is moving, the more energy it possesses.</td><td>Energy and Matter
• Energy can be transferred in various ways and between objects.</td></tr>
<tr><td>Clarifications and Content Limits</td><td colspan="3">Content Limits
• Assessment does not include quantitative measures of changes in the speed of an object or on any precise or quantitative definition of energy.
• Students are expected to know that energy can be expressed through sound, heat, light, and motion.
• <u>Students do not need to know:</u> Students do not need to know how to calculate speed, the change in speed (acceleration), or energy. This standard is limited to making strictly qualitative or comparative observations.</td></tr>
<tr><td>Science Vocabulary Students Are Expected to Know</td><td colspan="3">Volume, collision, heat transfer, spring (coil), forms of energy (sound, heat, light, motion), conservation of energy, stored energy, energy transfer, gravity.</td></tr>
<tr><td>Science Vocabulary Students Are Not Expected to Know</td><td colspan="3">Potential energy, kinetic energy, thermal energy, acceleration, velocity.</td></tr>
<tr><td colspan="4">Phenomena</td></tr>
<tr><td>Context/ Phenomena</td><td colspan="3">Some example phenomena for 4-PS3-1:
• One drum can be used to produce loud or quiet percussion sounds.
• A small bouncing basketball sounds louder than a large bouncing basketball.
• Damage caused during a high-speed collision is greater than when speeds are slower.
• A ceramic bowl dropped from a greater height will have a larger debris pattern.</td></tr>
<tr><td colspan="4">This Performance Expectation and associated Evidence Statements support the following Task Demands.</td></tr>
<tr><td colspan="4">Task Demands</td></tr>
<tr><td colspan="4">1. Articulate, describe, illustrate, or select the relationships, interactions, and/or processes to be explained. This may entail sorting relevant from irrelevant information or features.**</td></tr>
<tr><td colspan="4">2. Express or complete a causal chain explaining that changes in energy and speed are related. This may include indicating directions of causality in an incomplete model such as a flow chart or diagram, or completing cause-and-effect chains.*</td></tr>
<tr><td colspan="4">3. Identify evidence supporting the inference of causation that is expressed in a causal chain.</td></tr>
<tr><td colspan="4">4. Use an explanation to predict how the speed of an object changes given a change in energy or how the expression of energy will change given a change in speed.</td></tr>
<tr><td colspan="4">5. Describe, identify, and/or select information needed to support an explanation.</td></tr>
</table>

*denotes those task demands which are deemed appropriate for use in stand-alone item development

**TD 1 should only be used if paired with TD2. TD 2 can be used alone.

This page intentionally left blank.

<table>
<tr><td>Performance Expectation</td><td colspan="3">4-PS3-2
Make observations to provide evidence that energy can be transferred from place to place by sound, light, heat, and electric currents.</td></tr>
<tr><td>Dimensions</td><td>Planning and Carrying Out Investigations
• Make observations to produce data to serve as the basis for evidence for an explanation of a phenomena or to test a design solution.</td><td>PS3.A: Definitions of Energy
• Energy can be moved from place to place by moving objects — or through sound, light, or electric currents.

PS3.B: Conservation of Energy and Energy Transfer
• Energy is present whenever there are moving objects, sound, light, or heat. When objects collide, energy can be transferred from one object to another, thereby changing their motion. In such collisions, some energy is typically also transferred to the surrounding air; as a result, the air gets heated and sound is produced.
• Light also transfers energy from place to place.
• Energy can also be transferred from place to place by electric currents, which can then be used locally to produce motion, sound, heat, or light. The currents may have been produced to begin with by transforming the energy of motion into electrical energy.</td><td>Energy and Matter
• Energy can be transferred in various ways and between objects.</td></tr>
<tr><td>Clarifications and Content Limits</td><td colspan="3">Content Limits
• Assessment does not include quantitative measurements of energy.
• Identifying how energy is transferred (example: conduction vs. convection) is not part of this PE.
• <u>Students do not need to know</u>: Students do not need to know how to do energy calculations. This standard is limited to strictly making observations. Students should know that energy can be given off as heat or light, but not specifics such as convection, thermal radiation, etc.</td></tr>
<tr><td>Science Vocabulary Students Are Expected to Know</td><td colspan="3">Collision, speed, flow, heat conduction, conversion.</td></tr>
<tr><td>Science Vocabulary Students Are Not Expected to Know</td><td colspan="3">Kinetic energy, potential energy, radiation, convection, transmission, reflection, decibels, resonance, friction, hertz, electromagnetic radiation, magnitude, motion energy, electric circuit, thermal, conservation of energy.</td></tr>
<tr><td colspan="4">Phenomena</td></tr>
<tr><td>Context/ Phenomena</td><td colspan="3">Some example phenomena for 4-PS3-2:
• A light bulb can be powered using the motion of a hamster wheel.
• A drinking glass can be broken by a person singing a certain note.
• A fan (with blades angled at 45 degrees) will spin when placed safely over burning candles.
• Touching a Van der Graaf generator will make your hair stick up.</td></tr>
<tr><td colspan="4">This Performance Expectation and associated Evidence Statements support the following Task Demands.</td></tr>
<tr><td colspan="4">Task Demands</td></tr>
</table>

1. Identify the materials/tools needed for an investigation of how energy is transferred from place to place through heat, sound, light, or electric currents.
2. Identify the data that should be collected in an investigation of how energy is transferred from one place to another through heat, sound, light, or electric currents.
3. Make and/or record observations about the transfer of energy from one place to another via heat, sound, light, or electric currents.**
4. Interpret and/or communicate the data from an investigation.**
5. Select, describe, or illustrate a prediction made by applying the findings from an investigation.

*denotes those task demands which are deemed appropriate for use in stand-alone item development
**TD3 and TD4 must be used together.

<table>
<tr><td>Performance Expectation</td><td colspan="3">4-PS3-3
Ask questions and predict outcomes about the changes in energy that occur when objects collide.</td></tr>
<tr><td>Dimensions</td><td>Asking Questions and Defining Problems
• Ask questions that can be investigated and predict reasonable outcomes based on patterns such as cause-and-effect relationships.</td><td>PS3.A: Definitions of Energy
• Energy can be moved from place to place by moving objects or through sound, light, or electric currents.

PS3.B: Conservation of Energy and Energy Transfer
• Energy is present whenever there are moving objects, sound, light, or heat. When objects collide, energy can be transferred from one object to another, thereby changing their motion. In such collisions, some energy is typically also transferred to the surrounding air; as a result, the air gets heated and sound is produced.

PS3.C: Relationship Between Energy and Forces
• When objects collide, the contact forces transfer energy so as to change the objects’ motions.</td><td>Energy and Matter
• Energy can be transferred in various ways and between objects.</td></tr>
<tr><td>Clarifications and Content Limits</td><td colspan="3">Clarification Statements
• Emphasis is on the change in the energy due to the change in speed, not on the forces, as objects interact.

Content Limits
• Assessment does not include quantitative measurements of energy.
• <u>Students do not need to know</u>: names of energy types, how to calculate energy or forces</td></tr>
<tr><td>Science Vocabulary Students Are Expected to Know</td><td colspan="3">Electric currents, speed, flow, conversion, motion, magnets, magnetism, heat conduction.</td></tr>
<tr><td>Science Vocabulary Students Are Not Expected to Know</td><td colspan="3">Kinetic energy, potential energy, friction, force fields, vector, magnitude, elastic, inelastic.</td></tr>
<tr><td colspan="4">Phenomena</td></tr>
<tr><td>Context/ Phenomena</td><td colspan="3">Some example phenomena for 4-PS3-3:
• A large wave crashes into the cliffs of Étretat and some rocks are knocked loose. A small wave then crashes into the cliffs.
• A person hits a nail with a hammer and the nail is driven into a board. The person swings the hammer again, but misses the nail.
• A person walks down a hallway. The sound of their shoes on the floor can be heard many feet away. The person then runs down the hallway.
• A bowler rolls a ball down a lane. It slams into the pins and knocks several of them down. After the pins are reset, the bowler rolls the ball down the lane again. The ball misses and knocks down no pins.</td></tr>
<tr><td colspan="4">This Performance Expectation and associated Evidence Statements support the following Task Demands.</td></tr>
<tr><td colspan="4">Task Demands</td></tr>
</table>

1. Select or identify from a collection, including distractors, questions that will help clarify the properties that are correlated with the changes in energy that occur in the phenomenon. In addition to distractors that are plausible responses, distractors may include non-testable ("nonscientific") questions.
2. Identify, describe, or select from a collection, including distractors, characteristics to be manipulated or held constant while gathering information to answer a well-articulated question.
3. Select or describe conclusions relevant to the question posed and supported by the data, especially conclusions about causes and effects.
4. Predict outcomes when properties or proximity of the objects are changed, given the inferred cause-and-effect relationships.
5. Describe, identify, gather, and/or select information needed to identify patterns that can be used to predict outcomes about the changes in energy.

Performance Expectation	**4-PS3-4** Apply scientific ideas to design, test, and refine a device that converts energy from one form to another.		
Dimensions	**Constructing Explanations and Designing Solutions** • Apply scientific ideas to solve design problems.	**PS3.B: Conservation of Energy and Energy Transfer** • Energy can also be transferred from place to place by electric currents, which can then be used locally to produce motion, sound, heat, or light. The currents may have been produced to begin with by transforming the energy of motion into electrical energy. **PS3.D: Energy in Chemical Processes and Everyday Life** • The expression "produce energy" typically refers to the conversion of stored energy into a desired form for practical use. **ETS1.A: Defining Engineering Problems** • Possible solutions to a problem are limited by available materials and resources (constraints). The success of a designed solution is determined by considering the desired features of a solution (criteria). Different proposals for solutions can be compared on the basis of how well each one meets the specified criteria for success or how well each takes the constraints into account.	**Energy and Matter** • Energy can be transferred in various ways and between objects.
Clarifications and Content Limits	**Clarification Statements** • Examples of devices could include electric circuits that convert electrical energy into motion energy of a vehicle, light, or sound; and, a passive solar heater that converts light into heat. Examples of constraints could include the materials, cost, or time to design the device. **Content Limits** • Devices should be limited to those that convert motion energy to electric energy or use stored energy to cause motion or produce light or sound.		
Science Vocabulary Students Are Expected to Know	Magnetic, motion, speed, conservation, gravitational, battery, conversion, properties, chemical.		
Science Vocabulary Students Are Not Expected to Know	Mass, net force, velocity, relative position, constant speed, direction of motion, direction of a force, deceleration, independent, economic, control, impact, inertia, Newton's laws (1st, 2nd, 3rd), stationary, frame of reference, potential energy, mechanical energy, kinetic energy, conserve, relative, chemical energy.		
Phenomena			
Context/ Phenomena	Engineering practices are built around meaningful design problems rather than phenomena. For this performance expectation, a design problem and associated competing solutions will replace phenomena. Some examples of design problems for 4-PS3-4:		

<table>
<tr><td></td><td>A front door does not have an alarm. Any alarm that is added needs to be heard in the back hallway.
A person hiking on a hot day needs to take a fan to stay cool. The fan must be small so that it does not add to the weight of the hiker’s pack but must also last the entire hike.
The water in a house is heated with electricity purchased from a power company. A decision is made to instead heat the water using electricity generated with solar panels on the roof. The water heater must heat enough water to meet the needs of the home but the cost of installation and/or maintenance cannot exceed the family’s budget.
A motor is added to a toy car for a race. The motor must be able to move the car across a room at a high speed.</td></tr>
<tr><td colspan="2">This Performance Expectation and associated Evidence Statements support the following Task Demands.</td></tr>
<tr><td colspan="2">Task Demands</td></tr>
<tr><td colspan="2">1. Express or complete a causal chain explaining how energy can be transferred via electric current to produce light, sound, heat, and/or motion. This may include indicating directions of causality in an incomplete model such as a flow chart or diagram, or completing cause-and-effect chains.</td></tr>
<tr><td colspan="2">2. Identify evidence supporting the inference of causation that is expressed in a causal chain.</td></tr>
<tr><td colspan="2">3. Use an explanation to predict how the motion, sound, heat, or light of an object changes, given a change in electrical energy—or, how the expression of energy will change, given a change in the conversion of stored energy.</td></tr>
<tr><td colspan="2">4. Identify or assemble from a collection, including distractors, the relevant aspects of the problem that given design solutions, if implemented, will resolve/improve. The design solution must convert energy from one form to another within the content limits.</td></tr>
<tr><td colspan="2">5. Using given information, select or identify constraints that the device that converts energy from one form to another must meet OR criteria against which it should be judged.</td></tr>
<tr><td colspan="2">6. Using given information, design, propose, illustrate, assemble, test, or refine a potential device (prototype) that converts energy from one form to another.</td></tr>
</table>

<table>
<tr><td>Performance Expectation</td><td colspan="3">4-PS4-1
Develop a model of waves to describe patterns in terms of amplitude and wavelength, and that waves can cause objects to move.</td></tr>
<tr><td>Dimensions</td><td>Developing and Using Models
• Develop a model using an analogy, example, or abstract representation to describe a scientific principle.</td><td>PS4.A: Wave Properties
▪ Waves, which are regular patterns of motion, can be made in water by disturbing the surface. When waves move across the surface of deep water, the water goes up and down in place; there is no net motion in the direction of the wave except when the water meets a beach.
▪ Waves of the same type can differ in amplitude (height of the wave) and wavelength (spacing between wave peaks).</td><td>Patterns
• Similarities and differences in patterns can be used to sort, classify, and analyze simple rates of change for natural phenomena.</td></tr>
<tr><td>Clarifications and Content Limits</td><td colspan="3">Clarification Statements
• Examples of models could include diagrams, analogies, and physical models using wire to illustrate wavelength and amplitude of waves.
• Acceptable clusters may include: amplitude and wavelength, motion of an object, or both.

Content Limits
• Limited to physically visible mechanical waves.
• Assessment does not include interference effects, electromagnetic waves, non-periodic waves, or quantitative models of amplitude and wavelength.
• Examples of objects being moved by waves are limited to up and down motion. Horizontal motion is above grade level due to the other factors involved.
• Don’t directly reference energy. Energy is addressed in 4-PS3.
• <u>Students do not need to know</u>:
o Types of waves: sound, light, non-periodic, compression
o Particle movement
o Quantitative models
o Behaviors of waves: absorption, reflection, refraction, transmission, interactions with different materials (angle of incidence, amount of reflection or absorption, light being refracted into colors). Reflection is limited to the concept. How waves are reflected and the details of reflection (as well as other behaviors) are covered in MS-PS4-2.
o Wave calculations
o Motion of objects in the ocean due to ocean currents</td></tr>
<tr><td>Science Vocabulary Students Are Expected to Know</td><td colspan="3">Crest, trough, peak, rate, property, medium, period</td></tr>
<tr><td>Science Vocabulary Students Are Not Expected to Know</td><td colspan="3">Electromagnetic, compression, particle, transmission, seismic wave, radio wave, microwave, infrared, ultraviolet, gamma rays, x-rays, angle of incidence, concave, convex, diffraction, constructive interference, destructive interference, resonance, refraction, absorption, reflection, pitch, sound wave, light wave.</td></tr>
<tr><td colspan="4" align="center">Phenomena</td></tr>
<tr><td>Context/</td><td colspan="3">Some example phenomena for 4-PS4-1:</td></tr>
</table>

Phenomena	• A boat floating in the ocean is tied to a pier. The boat rises and falls with the waves. • Two students hold ends of a rope. One student lifts her end, and then drops it toward the ground. The rope forms a wave that travels from that student to the other student. • The sand waves on a windy beach get bigger and more pronounced over time. They are regular and evenly spaced. • A surfer riding a wave stays up if she moves along the wave but falls as soon as she stops moving.

This Performance Expectation and associated Evidence Statements support the following Task Demands.
Task Demands
1. Select or identify the components of a model that are needed to describe wave behavior, patterns of wave creation, and/or the motion of objects carried on/by waves. Components might include the source, amplitude, frequency, and/or wavelength.
2. Manipulate the components of a model to demonstrate properties, processes, and/or events that result in the patterns of wave behavior that are identified in the phenomenon. These patterns of wave behavior can include creation and replication of waves.
3. Describe, select, or identify the relationships among components of a model that describe wave behavior, patterns of wave creation, and/or the motion of objects carried on or by a wave.
4. Given a model of waves, illustrate the way in which the wave changes to yield a given result (more movement, less movement) and/or identify the result based on changes to the wave.
5. Make predictions about the effects of changes in model components (e.g., energy of wave source, distance from wave source), the amplitude or wavelength of a wave, or motion of objects affected by the wave. Item writer: Do not directly reference the energy of the wave source. Instead, show the speed and size of the object causing the wave, etc.

<table>
<tr><td>Performance Expectation</td><td colspan="3">4-PS4-2
Develop a model to describe that light reflecting from objects and entering the eye allows objects to be seen.</td></tr>
<tr><td>Dimensions</td><td>Developing and Using Models
• Develop a model to describe phenomena.</td><td>PS4.B: Electromagnetic Radiation
• An object can be seen when light reflected from its surface enters the eyes.</td><td>Cause and Effect
• Cause-and-effect relationships are routinely identified.</td></tr>
<tr><td>Clarifications and Content Limits</td><td colspan="3">Content Limits
• Assessment does not include:
○ knowledge of specific colors reflected and seen;
○ the cellular mechanisms of vision;
○ how the retina works.</td></tr>
<tr><td>Science Vocabulary Students Are Expected to Know</td><td colspan="3">Energy, light ray, reflection, reflective, surface</td></tr>
<tr><td>Science Vocabulary Students Are Not Expected to Know</td><td colspan="3">Particle, transmission, angle of incidence, angle of reflection, concave, convex, diffraction, constructive interference, destructive interference, refraction, absorption, wave, field, illuminate, diffuse reflection, specular reflection, spectrum, prism.</td></tr>
<tr><td colspan="4">Phenomena</td></tr>
<tr><td>Context/ Phenomena</td><td colspan="3">Some example phenomena for 4-PS4-2:
• A person can see a cat in the mirror. The cat is otherwise hidden from view.
• A performance is being watched by a person. Another person stands up and blocks the view.
• A flashlight is pointed at a door in a dark room. The door is the only object seen in the room.
• The surface of a lake is very still. The reflection of a tree on the bank can be seen on the lake’s surface.</td></tr>
<tr><td colspan="4">This Performance Expectation and associated Evidence Statements support the following Task Demands.</td></tr>
<tr><td colspan="4">Task Demands</td></tr>
<tr><td colspan="4">1. Identify the components needed to model the phenomenon. Components might include the light, the light source, the object, the path the light follows, and the eye.</td></tr>
<tr><td colspan="4">2. Complete an illustration or flow chart that is capable of representing how light reflecting from objects and entering the eye allows objects to be seen. This <u>does not</u> include labeling an existing diagram.</td></tr>
<tr><td colspan="4">3. Manipulate the components of a model to demonstrate the changes, properties, processes, and/or events that act to result in the phenomenon.</td></tr>
<tr><td colspan="4">4. Make predictions about the effects of changes in the model, particularly using mirrors, changing positions of light sources, objects, and the eye. Predictions can be made by manipulating model components, completing illustrations, or selecting from lists with distractors.</td></tr>
<tr><td colspan="4">5. Identify missing components, relationships, or other limitations of the model.</td></tr>
<tr><td colspan="4">6. Describe, select, or identify the relationships among components of a model that describe how light reflecting from objects and entering the eye allows objects to be seen.</td></tr>
</table>

This page intentionally left blank.

<table>
<tr><td>Performance Expectation</td><td colspan="3">4-PS4-3
Generate and compare multiple solutions that use patterns to transfer information.</td></tr>
<tr><td>Dimensions</td><td>Constructing Explanations and Designing Solutions
• Generate and compare multiple solutions to a problem, based on how well they meet the criteria and constraints of the design solution.</td><td>PS4.C: Information Technologies and Instrumentation
• Digitized information can be transmitted over long distances without significant degradation. High-tech devices, such as computers or cell phones, can receive and decode information—convert it from digitized form to voice—and vice versa.

ETS1.C: Optimizing the Design Solution
• Different solutions need to be tested in order to determine which of them best solves the problem, given the criteria and the constraints.</td><td>Patterns
• Similarities and differences in patterns can be used to sort and classify designed products.</td></tr>
<tr><td>Clarifications and Content Limits</td><td colspan="3">Clarification Statements
• Examples of solutions could include:
  • drums sending coded information through sound waves;
  • using a grid of 1’s and 0’s, representing black and white, to send information about a picture;
  • using Morse code to send text.

Content Limits
• <u>Students do not need to know:</u>
  o the different parts of the electromagnetic spectrum (visible, microwave, x-ray, radio wave, etc.);
  o binary coding or how it works;
  o that light is made up of an electric and magnetic field;
  o transverse vs. longitudinal waves;
  o how information gets encoded;
  o how different forms of communicating information work (Morse code vs. something like a telephone).</td></tr>
<tr><td>Science Vocabulary Students Are Expected to Know</td><td colspan="3">Amplitude, wavelength, reflect, vibrate, vibration, absorb, properties, sound wave, wave, communicate, electricity, coded, Morse code, digital, store, transfer, convert.</td></tr>
<tr><td>Science Vocabulary Students Are Not Expected to Know</td><td colspan="3">Light emission, light refraction, transmit, wave peaks, light wave, electromagnetic, frequency, radiation, wave packet, light scattering, light transmission, electric field, magnetic field, photon, radio wave, x-ray, binary, electron, pixel, CCD, transverse, longitudinal.</td></tr>
<tr><td colspan="4" align="center">Phenomena</td></tr>
<tr><td>Context/ Phenomena</td><td colspan="3">Some example phenomena for 4-PS4-3:
• In July 2015, the New Horizons Space Probe flew past Pluto. The space probe is tasked with taking detailed pictures of Pluto so that scientists on Earth can study its features. However, the spacecraft can only send sequences of numbers back to Earth.</td></tr>
</table>

	• A man wants to send an urgent message to his wife who is a long distance away. It would take too long to drive to his wife and deliver the message himself. The only way he can communicate is through an electrical wire that is set up between the two locations. • Two people want to communicate a number 1 through 10 over a large distance. They have no telephones or other means of communication. They are close enough that they can see or hear each other, however, a river separates them so they cannot reach each other. • Two people want to communicate over a large distance. However, the power is out and so they cannot use the telephone. All they have is a string that is stretched between their two houses. Attached to the end of each string is a metal can. The messages they want to be able to send consists of numbers 1 through 10.

This Performance Expectation and associated Evidence Statements support the following Task Demands.
Task Demands
1. Articulate, describe, illustrate, or select the relationships, interactions, and/or processes to be explained. This may entail sorting relevant from irrelevant information or features.
2. Express or complete a causal chain explaining how each pattern is used to transmit information. This may include indicating directions of causality in an incomplete model such as a flow chart or diagram, or completing cause-and-effect chains.
3. Identify evidence supporting the inference of causation that is expressed in a causal chain.
4. Use an explanation to compare the two solutions and select which one is better for the transmitting of information.
5. Describe, identify, and/or select information needed to support an explanation.

<table>
<tr><td>Performance Expectation</td><td colspan="3">4-LS1-1
Construct an argument that plants and animals have internal and external structures that function to support survival, growth, behavior, and reproduction.</td></tr>
<tr><td>Dimensions</td><td>Engaging in Argument from Evidence
• Construct an argument with evidence, data, and/or a model.</td><td>LS1.A: Structure and Function
• Plants and animals have both internal and external structures that serve various functions in growth, survival, behavior, and reproduction.</td><td>Systems and System Models
• A system can be described in terms of its components and their interactions.</td></tr>
<tr><td>Clarifications and Content Limits</td><td colspan="3">Clarification Statements
• Examples of structures could include thorns, stems, roots, colored petals, heart, stomach, lung, brain, and skin.

Content Limits
• Assessment is limited to macroscopic structures within plant and animal systems.
• The student does not need to know about cellular structures like the nucleus, mitochondria, the Golgi apparatus or the endoplasmic reticulum.
• <u>The student does not need to know:</u> about organ systems like the circulatory system, reproductive system, or nervous system.</td></tr>
<tr><td>Science Vocabulary Students Are Expected to Know</td><td colspan="3">Brain, body, flow, flower, heart, lung, muscle, movement, grasp, habit, moisture, organization, petal, predator, prey, roots, skin, stem, stomach, temperature</td></tr>
<tr><td>Science Vocabulary Students Are Not Expected to Know</td><td colspan="3">Cell, detect, response, body plan, elastic, external, intellectual, internal, invertebrate, organ, vertebrate, multicellular, stimulus, tissue, enzyme, xylem, phloem, parenchyma, and cambium cells.</td></tr>
<tr><td colspan="4">Phenomena</td></tr>
<tr><td>Context/ Phenomena</td><td colspan="3">Some example phenomena for 4-LS1-1:
• In a field of grass, a butterfly lands on one of the only red poppy flowers in sight.
• A manta ray has a flat circular body. Its fins spread out like wings from its body.
• A pelican can hold up to 3 gallons of water in its pouch.
• A student sees a hollow, brown copy of a cicada insect attached to the bark of a tree.</td></tr>
<tr><td colspan="4">This Performance Expectation and associated Evidence Statements support the following Task Demands.</td></tr>
<tr><td colspan="4">Task Demands</td></tr>
<tr><td colspan="4">1. Identify evidence or patterns in the data that support inferences and/or determine relationships between a particular structure of an organism and a function that supports survival, growth, behavior, and reproduction.</td></tr>
<tr><td colspan="4">2. Understand and generate simple bar graphs or tables to document patterns, trends, or relationships between a particular structure of an organism and a function that supports survival, growth, behavior, and reproduction.</td></tr>
</table>

3. Sort observations/evidence into those that appear to support or not support an argument.
4. Based on the provided data, identify or describe a claim regarding the relationship between a structure of an organism and a function that supports survival, growth, behavior, and reproduction.
5. Summarize or organize given data or other information to support or refute a claim regarding an organism's structure and its function.
6. Sort, tabulate, classify, separate, and/or categorize relevant from irrelevant information regarding an organism's structure and its function.

<table>
<tr><td>Performance Expectation</td><td colspan="3">4-LS1-2
Use a model to describe that animals receive different types of information through their senses, process the information in their brain, and respond to the information in different ways.</td></tr>
<tr><td>Dimensions</td><td>Developing and Using Models
• Use a model to test interactions concerning the functioning of a natural system.</td><td>LS1.D: Information Processing
• Different sense receptors are specialized for particular kinds of information, which may be then processed by the animal’s brain. Animals are able to use their perceptions and memories to guide their actions.</td><td>Systems and System Models
• A system can be described in terms of its components and their interactions.</td></tr>
<tr><td>Clarifications and Content Limits</td><td colspan="3">Clarification Statements
• Emphasis is on systems of information transfer.

Content Limits
• Assessment does not include the mechanisms by which the brain stores and recalls information or the mechanisms of how sensory receptors function.</td></tr>
<tr><td>Science Vocabulary Students Are Expected to Know</td><td colspan="3">Lens, vision, hearing, muscle, ear, middle ear, outer ear, inner ear, eardrum, response, habitat, eye, lens, memory</td></tr>
<tr><td>Science Vocabulary Students Are Not Expected to Know</td><td colspan="3">Sensory, brain, cells, retina, pupil, saliva, salivary gland, vibration, cornea, iris, brainstem, consumer, nerve, optic nerve, nerve cell, nerve tissue, nerve impulse, connecting nerve, nerve fiber, organ system, reflex, reflex action, reaction time, cue.</td></tr>
<tr><td colspan="4">Phenomena</td></tr>
<tr><td>Context/ Phenomena</td><td colspan="3">Some example phenomena for 4-LS1-2:
• A bear cub in the woods cries out. Its mother immediately runs toward it.
• A deer walks in the woods. It turns suddenly and moves off in a different direction. A few minutes later, a skunk appears from the bushes.
• A cat sits on a stone wall. A mouse appears at the base of a nearby tree. The cat springs after the mouse.
• A hawk flies overhead. Suddenly, it dives toward the tall grass. A moment later, it returns to the sky, a snake in its claws.</td></tr>
<tr><td colspan="4">This Performance Expectation and associated Evidence Statements support the following Task Demands.</td></tr>
<tr><td colspan="4">Task Demands</td></tr>
<tr><td colspan="4">1. Select or identify from a collection of potential model components the components needed to model the phenomenon. Components might represent organ systems or parts of a system needed for collection and/or processing of sensory information.</td></tr>
<tr><td colspan="4">2. Assemble or complete, from a collection of potential model components, an illustration or flow chart that is capable of representing the flow and/or processing of sensory information in an animal. This <u>does not</u> include labeling an existing diagram.</td></tr>
</table>

3. Manipulate the components of a model to demonstrate the changes, properties, processes, and/or events that act to result in the phenomenon.*
4. Given models or diagrams of the flow and/or processing of sensory information in an animal, identify responses to sensory inputs and how they change in each scenario OR identify the properties of organs and/or organ systems that allow animals to respond to sensory information.*
5. Identify missing components, relationships, or other limitations of a model that shows the flow and/or processing of sensory information in an animal.
6. Describe, select, or identify the relationships among components of a model that describe how sensory information is processed or explain how an animal responds to sensory inputs.

*denotes those task demands which are deemed appropriate for use in stand-alone item development

<table>
<tr><td>Performance Expectation</td><td colspan="3">4-ESS1-1
Identify evidence from patterns in rock formations and fossils in rock layers to support an explanation for changes in a landscape over time.</td></tr>
<tr><td>Dimensions</td><td>Constructing Explanations and Designing Solutions
• Identify the evidence that supports particular points in an explanation.</td><td>ESS1.C: The History of Planet Earth
• Local, regional, and global patterns of rock formations reveal changes over time due to Earth forces, such as earthquakes. The presence and location of certain fossil types indicate the order in which rock layers were formed.</td><td>Patterns
• Patterns can be used as evidence to support an explanation.</td></tr>
<tr><td>Clarifications and Content Limits</td><td colspan="3">Clarification Statement
• Examples of evidence from patterns could include rock layers with marine shell fossils above rock layers with plant fossils and no shells, indicating a change from land to water over time, and a canyon with different rock layers in the walls and a river in the bottom, indicating that over time a river cut through the rock.

Content Limits
• Assessment does not include specific knowledge of the mechanism of rock formation or memorization of specific rock formations and layers.
• Assessment is limited to relative time.
• Assessment does not include earthquakes—the clarification statement focuses on geomorphology and landscape change through time. The focus is not on tectonics, despite its mention in the DCI.</td></tr>
<tr><td>Science Vocabulary Students are Expected to Know</td><td colspan="3">Weathering, erode, glacier, climate, fossil, landscape, shell, river, mountain, canyon, deposit, marine.</td></tr>
<tr><td>Science Vocabulary Students are Not Expected to Know</td><td colspan="3">Rock strata, ocean basins, glaciation, watersheds, geological, mountain chains, igneous rock, metamorphic rock, sedimentary rock, terrestrial, aquatic.</td></tr>
<tr><td colspan="4">Phenomena</td></tr>
<tr><td>Context/ Phenomena</td><td colspan="3">Sample phenomena for 4-ESS1-1:
• The rock walls on both sides of the Grand Canyon contain layers with marine fossils, interspersed with layers containing terrestrial fossils.
• Church Rock, New Mexico, is a very dry place far from the sea. However, exposures of rocks in the area contain many fossils of marine organisms.
• Axel Heiberg Island in the Canadian Arctic is too cold for trees to grow. However, sedimentary rocks on the island preserve hundreds of fossil stumps from large evergreen trees.
• Sihetun, China, is dry and mountainous. Sedimentary rocks exposed in the area preserve thousands of fish fossils. These sedimentary rocks are sandwiched between lava flow rocks. There are no active volcanoes in this part of China.</td></tr>
<tr><td colspan="4">This Performance Expectation and associated Evidence Statements support the following Task Demands.</td></tr>
<tr><td colspan="4">Task Demands</td></tr>
</table>

1. Describe, identify, and/or select evidence from patterns of rock formations and/or patterns of fossils in rock layers to support the explanations of changes in the landscape over time.
2. Express or complete a causal chain explaining changes in patterns of fossils in rock layers.
3. Identify patterns of rock formations and/or patterns of fossils in rock layers.

<table>
<tr><td>Performance Expectation</td><td colspan="3">4-ESS2-1
Make observations and/or measurements to provide evidence of the effects of weathering or the rate of erosion by water, ice, wind, or vegetation.</td></tr>
<tr><td>Dimensions</td><td>Planning and Carrying Out Investigations
• Make observations and/or measurements to produce data to serve as the basis for evidence for an explanation of a phenomenon.</td><td>ESS2.A: Earth Materials and Systems
• Rainfall helps to shape the land and affects the types of living things found in a region. Water, ice, wind, living organisms, and gravity break rocks, soils, and sediments into smaller particles and move them around.</td><td>Cause and Effect
• Cause and effect relationships are routinely identified, tested, and used to explain change.</td></tr>
<tr><td>Clarifications and Content Limits</td><td colspan="3">Clarification Statement
• Examples of variables to test could include: angle of slope in the downhill movement of water, amount of vegetation, speed of wind, relative rate of deposition, cycles of freezing and thawing of water, cycles of heating and cooling, and volume of water flow.

Content Limits
• Students aren’t expected to know the flow of energy that causes the phenomena.
• Assessment is limited to one form of erosion.
• Assessment does not include chemical erosion.
• <u>Students do not need to know</u>: Sedimentation, Earth’s interior, crystallization, minerals, the rock cycle, dynamic forces, feedback interactions, constructive forces, or deformation.</td></tr>
<tr><td>Science Vocabulary Students are Expected to Know</td><td colspan="3">Erosion, freeze, movement, cycle, weathering, ocean, sediment, vegetation, particle, earthquake, volcanoes, thaw.</td></tr>
<tr><td>Science Vocabulary Students are Not Expected to Know</td><td colspan="3">Composition, slope, continental boundaries, trench, minerals, plate tectonics, topography.</td></tr>
<tr><td colspan="4">Phenomena</td></tr>
<tr><td>Context/ Phenomena</td><td colspan="3">Some example phenomena for 4-ESS2-1:
• Rocks in the bottom of a river are usually smooth, but the rocks sitting on the ground nearby often have sharp edges and corners.
• Near its start in Colorado, the bed of the North Platte River is covered with boulders. Some five hundred miles away in Nebraska, the bed of the river is mostly sand.
• New gullies appear in a gravel driveway after a heavy rain.
• Over the course of a summer there is a series of major storms. At the end of the season, the channel of a small stream running through a grassy park is significantly wider than it was before the storms.</td></tr>
<tr><td colspan="4">This Performance Expectation and associated Evidence Statements support the following Task Demands.</td></tr>
<tr><td colspan="4">Task Demands</td></tr>
<tr><td colspan="4">1. Identify the factors that affect weathering or the rate of erosion by water, ice, wind, or vegetation.</td></tr>
</table>

2. Identify from a list the materials/tools needed for an investigation of how wind affects the factors that affect weathering or the rate of erosion by water, ice, wind, or vegetation.
3. Identify, among distractors, the outcome data that should be collected in the investigation.
4. Make and/or record observations about how input factors affect relevant outcomes while using fair tests in which variables are controlled.*
5. Make or communicate the conclusions from the investigation. Conclusions will be causal relationships.**

*denotes those task demands which are deemed appropriate for use in stand-alone item development
**TD5 can be used ONLY if used in concert with TD4

Performance Expectation	**4-ESS2-2** Analyze and interpret data from maps to describe patterns of Earth's features.		
Dimensions	**Analyzing and Interpreting Data** • Analyze and interpret data to make sense of phenomena using logical reasoning.	**ESS2.B: Plate Tectonics and Large-Scale System Interactions** • The locations of mountain ranges, deep ocean trenches, ocean floor structures, earthquakes, and volcanoes appear in patterns. Most earthquakes and volcanoes occur in bands that are often along the boundaries between continents and oceans. Major mountain chains form inside continents or near their edges. Maps can help locate the different land and water features areas of Earth.	**Patterns** • Patterns can be used as evidence to support an explanation.
Clarifications and Content Limits	**Clarification Statements** • Maps can include topographic maps of Earth's land and ocean floor, as well as maps of the locations of mountains, continental boundaries, volcanoes, and earthquakes. **Content Limits** • <u>Students do not need to know:</u> the tectonic processes that form Earth's features.		
Science Vocabulary Students Are Expected to Know	Earthquake, Earth's surface, crust, volcanic eruption, region, barrier, global, local, physical characteristic, ocean, force, landscape, mountain chain, mountain range, continental boundary, sea floor, collide, properties, ocean trench, pressure, topographic map.		
Science Vocabulary Students Are Not Expected to Know	Geologic, impact, magnitude, frequency, sediment deposition, ancient, ocean basin, rock layer movement, formation, continental shelf, deform, density, tectonic process, distribution, oceanic crust, plate boundary/collision, seafloor spreading.		
Phenomena			
Context/ Phenomena	For this performance expectation, the phenomena are the patterns of features on maps that the student examines. These patterns can sometimes be described with simple statements as shown below, but the actual phenomenon in each case is the pattern on the map. If descriptive statements are used, writers must be careful not to give the pattern or the point of the cluster away to the student. Some example phenomena for 4-ESS2-2: • There are active volcanoes in Alaska. There are no active volcanoes near Buffalo, New York. (If this statement were to be used to describe the map, then the students task would have to be something more than simply pointing out that there are volcanoes in Alaska and none near Buffalo, such as figuring out that Alaska is closer to a tectonic plate boundary than is New York.) • Earthquakes occur often in western South America. Earthquakes almost never occur on the eastern side of the continent. (If this statement were to be used to describe the map, then the student's task would have to be something more than simply pointing out that there are earthquakes on the eastern side more often than the western, such as figuring out that a plate boundary lies along the eastern coast of South America.) • Many volcanoes are found in a ring around the Pacific Ocean. There are fewer found on the edges of the Atlantic Ocean. (If this statement were to be used to describe the map, then		

	the students task would have to be something more than simply pointing out that there are many volcanoes around the Pacific and few around the Atlantic, such as figuring out that tectonic plate boundaries surround the Pacific Ocean.) • There are no mountain ranges in Kansas. There are many mountains in Washington State. (If this statement were to be used to describe the map, then the students task would have to be something more than simply pointing out that there are mountains in Washington and none in Kansas, such as figuring out that Washington is closer to a tectonic plate boundary than Kansas.)

This Performance Expectation and associated Evidence Statements support the following Task Demands.
Task Demands
1. Organize, arrange, or summarize map data and/or symbols to highlight/describe patterns of geological features on Earth's surface.**
2. Generate/construct graphs, tables, or assemblages of illustrations and/or labels, of map data that document patterns of geological features on Earth's surface. This may include sorting out distractors.*
3. Use relationships identified in the presented map data to predict the location of geological features on Earth's surface, such as mountain ranges, volcanoes, earthquake foci, and deep ocean trenches.*
4. Identify evidence or patterns in map data that support inferences about the patterns of geological features on Earth's surface.*

*denotes those task demands which are deemed appropriate for use in stand-alone item development
**TD1 may be used in combination with 2, 3, or 4 for stand-alone development.

<table>
<tr><td>Performance Expectation</td><td colspan="3">4-ESS3-1
Obtain and combine information to describe that energy and fuels are derived from natural resources and their uses affect the environment.</td></tr>
<tr><td>Dimensions</td><td>Obtaining, Evaluating, and Communicating Information
• Obtain and combine information from books and other reliable media to explain phenomena</td><td>ESS3.A: Natural Resources
• Energy and fuels that humans use are derived from natural sources, and their use affects the environment in multiple ways. Some resources are renewable over time, and others are not.</td><td>Cause and Effect
• Cause and effect relationships are routinely identified and used to explain change.</td></tr>
<tr><td>Clarifications and Content Limits</td><td colspan="3">Clarification Statements
• Examples of renewable energy resources could include:
  o Wind energy
  o Water behind dams
  o Sunlight
• Examples of non-renewable energy resources are:
  o fossil fuels
  o fissile materials
• Examples of environmental effects could include:
  o Loss of habitat due to dams
  o Loss of habitat due to surface mining
  o Air pollution from burning of fossil fuels

Content Limits
• The following things should be avoided:
  o Casting fossil fuels in a negative light and alternative fuels in a positive light
  o Pros and cons of one energy source vs. another
  o Negative effects of extracting and burning coal
  o Negative effects of fracking
  o Cause and effect of acid rain
  o The term “global warming”
  o <u>Students do not need to know:</u> how natural resources are used to generate energy (scientific specifics regarding how burning coal creates energy/how wind produces energy etc.).</td></tr>
<tr><td>Science Vocabulary Students are Expected to Know</td><td colspan="3">Recycle, reuse, coal, habitat, pollution, dam, population, atmosphere, oil, resource, fossil fuel, renewable, nonrenewable, conservation</td></tr>
<tr><td>Science Vocabulary Students are Not Expected to Know</td><td colspan="3">Agricultural, biosphere, mineral, geological, hydrothermal, metal ore, organic, deposition, petroleum, derive, extract, natural gas, oil shale, sustainability, tar sand</td></tr>
<tr><td colspan="4">Phenomena</td></tr>
<tr><td>Context/ Phenomena</td><td colspan="3">Some example phenomena for 4-ESS3-1
• A pipeline is built to transport oil from one location to another. As the oil moves across the landscape it leaks into a river along the way.</td></tr>
</table>

	• The Three Gorges dam was built along the Yangtze River in China to generate electricity. The Chinese dove tree lives along the Yangtze River. Building the dam affected this tree. • Several wind turbines are placed in a field to provide electricity to neighboring areas. To do this, forest land had to be cut down to provide space for the wind turbines. • Oil can be used to generate electricity. Oil can be found under the ocean. Seismic waves are used to locate the oil. Because of this, 100 melon head whales were displaced off the coast of Madagascar.

This Performance Expectation and associated Evidence Statements support the following Task Demands.
Task Demands
1. Organize and/or arrange (e.g., using illustrations and/or labels), or summarize data/information to highlight trends, patterns, or correlations.
2. Express or complete a causal chain explaining how energy and fuel that are derived from natural resources affect the environment. This may include indicating directions of causality in an incomplete model such as a flow chart or diagram, or completing cause and effect chains.*
3. Identify evidence supporting the inference of causation that is expressed in a causal chain.*
4. Identify patterns or evidence in the data that supports inferences about the effects that the usage of certain natural resources has on the environment.
5. Describe, identify, and/or Select information needed to support an explanation.

*denotes those task demands which are deemed appropriate for use in stand-alone item development

<table>
<tr><td>Performance Expectation</td><td colspan="3">4-ESS3-2
Generate and compare multiple solutions to reduce the impacts of natural Earth processes on humans.</td></tr>
<tr><td>Dimensions</td><td>Constructing Explanations and Designing Solutions
• Generate and compare multiple solutions to a problem based on how well they meet the criteria and constrains of the design solution</td><td>ESS3.B: Natural Hazards
• A variety of hazards result from natural processes (e.g., earthquakes, tsunamis, volcanic eruptions). Humans cannot eliminate the hazards but can take steps to reduce their impacts.

ETS1.B: Designing Solutions to Engineering Problems
• Testing a solution involves investigating how well it performs under a range of likely condition (secondary)</td><td>Cause and Effect
• Cause and effect relationships are routinely identified, tested, and used to explain change.</td></tr>
<tr><td>Clarifications and Content Limits</td><td colspan="3">Clarification Statements
• Examples of solutions could include designing an earthquake resistant building and improving monitoring of volcanic activity.

Content Limits
• Assessment is limited to earthquakes, floods, tsunamis, and volcanic eruptions.</td></tr>
<tr><td>Science Vocabulary Students are Expected to Know</td><td colspan="3">Environment, nature, recycle, reuse, coal, habitat, pollution, dam, population, atmosphere, oil, resource, fossil fuel, renewable, nonrenewable, conservation</td></tr>
<tr><td>Science Vocabulary Students are Not Expected to Know</td><td colspan="3">Agricultural, biosphere, mineral, geological, hydrothermal, metal ore, organic, deposition, petroleum, derive, extract, natural gas, oil shale, sustainability, tar sand</td></tr>
<tr><td colspan="4">Phenomena</td></tr>
<tr><td>Context/ Phenomena</td><td colspan="3">Engineering performance expectations are built around meaningful design problems rather than phenomena. In this case, the design problems involve reducing the impact of earthquakes, floods, tsunamis, and volcanic eruptions on humans. For this performance expectation, the design problem and competing solutions replace phenomena.

Example phenomena for 4-ESS3-2:
• Hurricanes generate high winds. Several building designs are being considered to construct buildings that could withstand the force of the wind.
• Eyjafjallajokull is an active volcano in Iceland. In preparation for future volcanic activity, several evacuation routes are being considered.</td></tr>
<tr><td colspan="4">This Performance Expectation and associated Evidence Statements support the following Task Demands.</td></tr>
<tr><td colspan="4">Task Demands</td></tr>
<tr><td colspan="4">1. Organize and/or arrange (e.g., using illustrations and/or labels), or summarize data/information to highlight trends, patterns, or correlations in data regarding human activity and natural hazards.</td></tr>
</table>

2.	Express or complete a causal chain explaining how humans can reduce the impact of natural hazards.
3.	Identify evidence supporting the inference of causation that is expressed in a causal chain.
4.	Identify patterns or evidence in the data that supports inferences about the ways humans can reduce impacts of natural hazards.
5.	Use an explanation to compare the two solutions and select which one is better for addressing the problem of the impact of natural hazards on humans and explain how well each solution meets the criteria and constraints of the design solution.
6.	Describe, select, or identify components of competing design solutions.

<table>
<tr><td>Performance Expectation</td><td colspan="3">5-PS1-1
Develop a model to describe that matter is made of particles too small to be seen.</td></tr>
<tr><td>Dimensions</td><td>Developing and Using Models
• Use models to describe phenomena.</td><td>PS1.A: Structure and Properties of Matter
• Matter of any type can be subdivided into particles that are too small to see, but even then the matter still exists and can be detected by other means. A model showing that gases are made from matter particles that are too small to see and are moving freely around in space can explain many observations, including the inflation and shape of a balloon and the effects of air on larger particles or objects.</td><td>Scale, Proportion, and Quantity
• Natural objects exist from the very small to the immensely large.</td></tr>
<tr><td>Clarifications and Content Limits</td><td colspan="3">Clarification Statements
• Examples of evidence supporting a model could include adding air to expand a basketball, compressing air in a syringe, dissolving sugar in water, and evaporating salt water.

Content Limits
• Assessment does not include the atomic-scale mechanism of evaporation and condensation or the defining of the unseen particles.
• Students are expected to know that matter can neither be destroyed nor created.</td></tr>
<tr><td>Science Vocabulary Students Are Expected to Know</td><td colspan="3">Substance, particle, solid, liquid, gas, vapor, steam, air, phase change, evaporate, boil, condense, freeze, melt, dissolve, mixture, chemical reaction, energy.</td></tr>
<tr><td>Science Vocabulary Students Are Not Expected to Know</td><td colspan="3">Atom, compound, molecule, chemical bond, solution, homogenous, heterogeneous, colloid, solute, solvent, precipitant, precipitate, reactant, product, air pressure, law of conservation of matter.</td></tr>
<tr><td colspan="4" align="center">Phenomena</td></tr>
<tr><td>Context/ Phenomena</td><td colspan="3">Some example phenomena for 5-PS1-1:
• A hissing sound can be heard as a bicycle wheel deflates.
• A sour odor can be smelled from milk that has been kept too long (or expired).
• When you pump air out of a closed bottle that is partially filled with marshmallows, the marshmallows expand in size. However, when you open the bottle, the marshmallows shrink back to their original size.
• When you place a lit match into a glass bottle and a boiled egg is set on the bottle opening, the egg eventually gets sucked into the bottle.</td></tr>
<tr><td colspan="4" align="center">This Performance Expectation and associated Evidence Statements support the following Task Demands.</td></tr>
<tr><td colspan="4" align="center">Task Demands</td></tr>
<tr><td colspan="4">1. Select or identify from a collection of potential model components, including distractors, the components needed to model the phenomenon. Components might include solid, liquid, or gas particles; particles of different substances; and representations of particle movement.</td></tr>
</table>

2. Assemble or complete — from a collection of potential model components — an illustration, flow chart, or causal chain that is capable of representing the particle nature of matter. This does not include labeling an existing diagram.
3. Manipulate the components of a model to demonstrate the changes, properties, processes, and/or events that act to result in the phenomenon.
4. Make predictions about the effects of changes in the movements of, distances between, or phases of the particles of matter under investigation. Predictions can be made by manipulating model components, completing illustrations, or selecting from lists with distractors.
5. Provided with models or diagrams of the particles of matter under investigation, identify the properties of the particles under investigation and how they change in each scenario. The properties of the particles may include the relative motions of, distances between, and phases of the particles.
6. Describe, select, or identify the relationships among components of a model that explains the observed effects of the particle nature of matter.

<table>
<tr><td>Performance Expectation</td><td colspan="3">5-PS1-2
Measure and graph quantities to provide evidence that regardless of the type of change that occurs when heating, cooling, or mixing substances, the total weight of matter is conserved.</td></tr>
<tr><td>Dimensions</td><td>Using Mathematics and Computational Thinking
• Measure and graph quantities such as weight to address scientific and engineering questions and problems.</td><td>PS1.A: Structure and Properties of Matter
• The amount (weight) of matter is conserved when it changes form, even in transitions in which it seems to vanish.

PS1.B: Chemical Reactions
• No matter what reaction or change in properties occurs, the total weight of the substances does not change.</td><td>Scale, Proportion, and Quantity
• Standard units are used to measure and describe physical quantities such as weight, time, temperature, and volume.</td></tr>
<tr><td>Clarifications and Content Limits</td><td colspan="3">Clarification Statements
• Examples of reactions or changes could include mixing, dissolving, and phase changes that form new substances.

Content Limits
• Assessment does not include distinguishing mass and weight.
• <u>Students do not need to know:</u> structure of atoms, specific chemical equations.</td></tr>
<tr><td>Science Vocabulary Students Are Expected to Know</td><td colspan="3">Weight, substance, matter, conservation, temperature, mixing, phase change, dissolving, properties, reaction, particles, gas, solid, liquid.</td></tr>
<tr><td>Science Vocabulary Students Are Not Expected to Know</td><td colspan="3">Mass, atoms, molecules, rates.</td></tr>
<tr><td colspan="4">Phenomena</td></tr>
<tr><td>Context/ Phenomena</td><td colspan="3">Some example phenomena for 5-PS1-2:
• A cup of water is taken out of the freezer and left on a counter. After some time, the frozen water melts.
• A cup of hot tea can dissolve more sugar than a cup of cold tea, but they both weigh the same after the mixing is complete.
• When mixed together, silver nitrate and sodium chloride forms a white solid that weighs the same as the individual silver nitrate and sodium chloride weighed.
• When water, baking soda, and calcium chloride are mixed inside a freezer bag, the bag gets hot and expands. The expanded freezer bag weighs the same as the ingredients did when they were separate.</td></tr>
<tr><td colspan="4">This Performance Expectation and associated Evidence Statements support the following Task Demands.</td></tr>
<tr><td colspan="4">Task Demands</td></tr>
<tr><td colspan="4">1. Make simple calculations using given data to calculate or estimate the total weight of a substance after heating, cooling, or mixing.</td></tr>
<tr><td colspan="4">2. Measure or graph data that can be used to calculate or estimate the total weight of a substance after heating, cooling, or mixing.</td></tr>
</table>

3. Describe and/or summarize data (e.g., using illustrations and/or labels) to identify/highlight trends, patterns, or correlations concerning the weight of the substances being investigated at the beginning and end of an investigation.
4. Compile and/or select, from given information, the particular data needed for a specific inference about the total weight of substances. This can include sorting out the relevant data from the overall body of given information.
5. Select, describe, or illustrate a prediction made by applying the findings from measurements or an investigation.
6. Use relationships identified in the data to explain that regardless of the type of change, the total weight of matter is conserved.

<table>
<tr><td>Performance Expectation</td><td colspan="3">5-PS1-3
Make observations and measurements to identify materials based on their properties.</td></tr>
<tr><td>Dimensions</td><td>Planning and Carrying Out Investigations
• Make observations and measurements to produce data to serve as the basis for evidence for an explanation of a phenomenon.</td><td>PS1.A: Structure and Properties of Matter
• Measurements of a variety of properties can be used to identify materials. (Boundary: At this grade level, mass and weight are not distinguished, and no attempt is made to define the unseen particles or explain the atomic-scale mechanism of evaporation and condensation.)</td><td>Scale, Proportion, and Quantity
• Standard units are used to measure and describe physical quantities such as weight, time, temperature, and volume.</td></tr>
<tr><td>Clarifications and Content Limits</td><td colspan="3">Clarification Statements
• Examples of materials to be identified could include baking soda and other powders, metals, minerals, and liquids.
• Examples of properties could include color, hardness, reflectivity, electrical conductivity, thermal conductivity, response to magnetic forces, and solubility.

Content Limits
• Assessment does not include density or distinguishing between mass and weight.
• <u>Students do not need to know:</u> chemical reaction equations, balancing reaction equations, atomic-level processes.</td></tr>
<tr><td>Science Vocabulary Students Are Expected to Know</td><td colspan="3">Electric, electrically charged, magnetic, magnetic attraction, conductor, change of state, substance, absorbency, evaporate, metal, vapor, conduction, relative, conservation of matter, phase change, dissolve, react, product</td></tr>
<tr><td>Science Vocabulary Students Are Not Expected to Know</td><td colspan="3">Insulator, element, reaction, boiling point, melting point, molecule, forms of matter, reactant, chemical compound, chemical reaction, atom</td></tr>
<tr><td colspan="4">Phenomena</td></tr>
<tr><td>Context/ Phenomena</td><td colspan="3">Some example phenomena for 5-PS1-3:
• Sugar and flour are white powdery substances. Sugar is soluble in water and flour is not.
• Three mineral crystals sit on a table. The three crystals are all the same color, resembling clear glass. However, they are all different minerals. One of them is quartz, one of them is halite, and the third is calcite.
• Two nails are on a table. When a magnet is placed over the nails, one of them moves from the table and sticks to the magnet.
• Two pieces of wood are hit with a hammer. One piece of wood has a depression/dent where the hammer hit it. The other does not have a dent/depression.</td></tr>
<tr><td colspan="4">This Performance Expectation and associated Evidence Statements support the following Task Demands.</td></tr>
<tr><td colspan="4">Task Demands</td></tr>
</table>

1. Identify from a list, including distractors, the materials or tools needed to observe or measure properties of matter to identify unknown materials.
2. Identify from a list, including distractors, the output data needed to identify or differentiate materials. **
3. Make and/or record observations or measurements from the investigation of the properties of materials.*
4. Interpret and/or communicate the data from the investigation of the properties of materials.
5. Make or communicate conclusions from the investigation of the properties of materials.

*denotes those task demands which are deemed appropriate for use in stand-alone item development
**TD2 may be used for stand-alone item if used with TD3

<table>
<tr><td>Performance Expectation</td><td colspan="3">5-PS1-4
Conduct an investigation to determine whether the mixing of two or more substances results in new substances.</td></tr>
<tr><td>Dimensions</td><td>Planning and Carrying Out Investigations
• Conduct an investigation collaboratively to produce data to serve as the basis for evidence, using fair tests in which variables are controlled and the number of trials are considered.</td><td>PS1.B: Chemical Reactions
• When two or more different substances are mixed, a new substance with different properties may be formed.</td><td>Cause and Effect
• Cause-and-effect relationships are routinely identified and used to explain change.</td></tr>
<tr><td>Clarifications and Content Limits</td><td colspan="3">Clarification Statements
• Students are not expected to be able to balance chemical equations, but should be able to complete simple mathematical (addition and subtraction) calculations in regard to starting materials and ending materials.

Content Limits
• Students are expected to know that matter is neither destroyed nor created.
• <u>Students do not need to know:</u> Chemical names, chemical symbols, general balanced equation {reactant + reactant → products}, and isotopes, specific chemical reaction types (e.g. oxidation, reduction, decomposition, and combustion).</td></tr>
<tr><td>Science Vocabulary Students Are Expected to Know</td><td colspan="3">Matter, substance, particle, chemical property, physical property, mass, volume, density, melting point, boiling point, freezing point, dissolve, flammable, odor, gas, solid, liquid, mixture, chemical reaction, gram(s), physical change, chemical change.</td></tr>
<tr><td>Science Vocabulary Students Are Not Expected to Know</td><td colspan="3">Reactant, product, atom, molecule, compound, chemical bond, law of conservation of mass, law of conservation of energy, intramolecular attractions, intermolecular attractions, solubility, solvent, solute, precipitant, rate of chemical reaction, acid, base, salt (as an ionic crystal), fusion, fission, homogeneous mixture, heterogeneous mixture, plasma, pH.</td></tr>
<tr><td colspan="4" align="center">Phenomena</td></tr>
<tr><td>Context/ Phenomena</td><td colspan="3">Some example phenomena for 5-PS1-4:
• A peach shrivels and becomes covered with mold.
• Over time, one metal changes color when exposed to rainwater. However, another metal exposed to rainwater does not.
• A bottle partially filled with vinegar sits on a counter. An empty balloon is partially filled with baking soda. When the open end of the balloon is stretched over the bottle top, a hissing/fizzing sound can be heard and the balloon expands.
• When sugar crystals are added to vinegar in a bowl, the crystals disappear. When crystals of baking soda are added to vinegar in a bowl, the mixture begins to bubble and foam.</td></tr>
<tr><td colspan="4" align="center">This Performance Expectation and associated Evidence Statements support the following Task Demands.</td></tr>
<tr><td colspan="4" align="center">Task Demands</td></tr>
<tr><td colspan="4">1. Identify from a list, including distractors, the properties that should be tested or the materials/tools needed in an investigation of the physical and chemical properties of the starting and ending substances involved in mixing.</td></tr>
</table>

2. Identify the outcome data that should be collected in an investigation of the physical and chemical properties of the starting and ending substances under investigation.
3. Make and/or record observations/data about the physical and chemical properties of the substances that are mixed and the substances resulting from the mixture.
4. Interpret and/or communicate the data from an investigation. This may include identifying/describing trends, patterns, or correlations among observations and data concerning the physical and chemical properties of the beginning and ending substances being investigated.
5. Explain or describe the causal processes that lead to the observed data.

<table>
<tr><td>Performance Expectation</td><td colspan="3">5-PS2-1
Support an argument that the gravitational force exerted by Earth on objects is directed down.</td></tr>
<tr><td>Dimensions</td><td>Engaging in Argument from Evidence
• Support an argument with evidence, data, or a model.</td><td>PS2.B: Types of Interactions
• The gravitational force of Earth acting on an object near Earth's surface pulls that object toward the planet's center.</td><td>Cause and Effect
• Cause and effect relationships are routinely identified and used to explain change.</td></tr>
<tr><td>Clarifications and Content Limits</td><td colspan="3">Clarification Statement
• "Down" is a local description of the direction that points toward the center of the spherical Earth.

Content Limits
• Assessment does not include mathematical representation of gravitational force.
• Study of gravity is limited to gravity on Earth.
• <u>Students do not need to know</u>: Calculations for weight (weight = mass • gravity), free fall, terminal velocity, weightlessness, air resistance, friction, black holes, inertia, Newton's law of universal gravitation, vacuum.</td></tr>
<tr><td>Science Vocabulary Students are Expected to Know</td><td colspan="3">Sun, gravity, space, flow, magnet, period (time), charge, Earth's rotation, solar system, spherical, exert, transfer, mass, orbital, mass, volume</td></tr>
<tr><td>Science Vocabulary Students are Not Expected to Know</td><td colspan="3">Attractive, direction of force, direction of motion, field, linear, nonlinear, gravitational energy, gravitational field, magnetic field, permeate.</td></tr>
<tr><td colspan="4">Phenomena</td></tr>
<tr><td>Context/ Phenomena</td><td colspan="3">Sample phenomena for 5-PS2-1:
• A hard rubber ball dropped in a pool falls more slowly than the same ball dropped on land.
• A feather released on top of a cliff on a breezy day seems to fly away, while a similar feather dropped on flat ground on a breezy day lands on the ground.
• A small piece of clay set on the top of a globe stays in place, but when you put it on the bottom of the globe it drops off. A piece of clay put at the real north pole stays in place, and also stays in place on the real south pole.
• A basketball flies in an arc before going through the basket.</td></tr>
<tr><td colspan="4">This Performance Expectation and associated Evidence Statements support the following Task Demands.</td></tr>
<tr><td colspan="4">Task Demands</td></tr>
<tr><td colspan="4">1. Sort observations into those that appear to support competing (given) arguments, or into those that support, contradict, or are not relevant to a given argument. Observations are from animations, simulations, or other given material.</td></tr>
<tr><td colspan="4">2. Sort, tabulate, classify, separate, and/ or categorize relevant from irrelevant evidence (observations) or data.</td></tr>
<tr><td colspan="4">3. Select from a given collection additional relevant observations that would help distinguish between competing arguments or the veracity of a single argument.</td></tr>
</table>

4. Select, identify, or describe apparent counterexamples to a supported argument.
5. Identify from a given collection or explain in writing flaws in observation that lead to an apparent counterexample, or explain the counterexample in terms of grade-level appropriate properties gravity, or other simple forces from earlier grade levels.
6. Sort statements into categories such as facts, reasonable judgments based on available facts, and speculation.
7. Clearly articulate the evidence supporting and contradicting an argument, noting how the evidence supports or contradicts the argument (hand scored).*
8. Predict outcomes when properties or proximity of the objects are changed, given the inferred cause and effect relationships. This can be done by describing outcomes, or selecting or identifying outcomes from lists.

*denotes those task demands which are deemed appropriate for use in stand-alone item development

<table>
<tr><td>Performance Expectation</td><td colspan="3">5-PS3-1
Use models to describe that energy in animals' food (used for body repair, growth, motion, and to maintain body warmth) was once energy from the sun.</td></tr>
<tr><td>Dimensions</td><td>Developing and Using Models
• Use models to describe phenomena.</td><td>PS3.D: Energy in Chemical Processes and Everyday Life
• The energy released [from] food was once energy from the sun that was captured by plants in the chemical process that forms plant matter.

LS1.C: Organization for Matter and Energy Flow in Organisms
• Food provides animals with the materials they need for body repair and growth and the energy they need to maintain body warmth and for motion. (secondary)</td><td>Energy and Matter
• Energy can be transferred in various ways and between objects.</td></tr>
<tr><td>Clarifications and Content Limits</td><td colspan="3">Clarification Statements
• Examples of models could include diagrams and flow charts.

Content Limits
• Assessment does not include photosynthesis.
• <u>Students do not need to know:</u> photosynthesis equation</td></tr>
<tr><td>Science Vocabulary Students are Expected to Know</td><td colspan="3">Energy, matter, transfer, light</td></tr>
<tr><td>Science Vocabulary Students are Not Expected to Know</td><td colspan="3">Photosynthesis, metabolism, atoms, chemicals, reaction, radiation</td></tr>
<tr><td colspan="4">Phenomena</td></tr>
<tr><td>Context/ Phenomena</td><td colspan="3">Some example phenomena for 5-PS3-1:
• Cows eat grass that grew in the sun.
• Termites eat the wood in trees.
• Caterpillars eat leaves and grow big.
• Koalas mainly eat eucalyptus leaves.</td></tr>
<tr><td colspan="4">This Performance Expectation and associated Evidence Statements support the following Task Demands.</td></tr>
<tr><td colspan="4">Task Demands</td></tr>
<tr><td colspan="4">1. Select or identify, from a collection of potential model components, including distractors, the parts of a model need to describe the flow of energy among plants, animals, and the sun.</td></tr>
<tr><td colspan="4">2. Assemble or complete a model representing the flow of energy among plants, animals, and the sun.</td></tr>
<tr><td colspan="4">3. Manipulate the components of a model to demonstrate properties, processes, and/or events that result in the flow of energy among plants, animals, and the sun, including the relationships of organisms and/or the cycles of energy and/or matter.</td></tr>
</table>

4. Articulate, describe, illustrate, select, or identify the relationships among components of a model that describe the movement of matter among plants, animals, and the sun.
5. Make predictions about the effects of changes in model components including the substitution, elimination, or addition of energy and/or an organism and the result.

This page intentionally left blank.

<table>
<tr><td>Performance Expectation</td><td colspan="3">5-LS1-1
Support an argument that plants get the materials they need for growth chiefly from air and water.</td></tr>
<tr><td>Dimensions</td><td>Engaging in Argument from Evidence
• Support an argument with evidence, data, or model.</td><td>LS1.C: Organization for Matter and Energy Flow in Organisms
• Plants acquire their material for growth chiefly from air and water.</td><td>Energy and Matter
• Matter is transported into, out of, and within systems.</td></tr>
<tr><td>Clarifications and Content Limits</td><td colspan="3">Clarification Statements
• Emphasis is on the idea that plant matter comes mostly from air and water, not from the soil.

Content Limits
• Assessment does not include photosynthesis or the photosynthesis reaction equation. Students should know that plants carry out photosynthesis for energy, but they do not need to know the specifics of the process or equation.</td></tr>
<tr><td>Science Vocabulary Students Are Expected to Know</td><td colspan="3">Organism, algae, atmosphere, consumer, cycle, matter, product, transport, chemical, convert, transfer, energy flow, flow chart, conservation, nutrients.</td></tr>
<tr><td>Science Vocabulary Students Are Not Expected to Know</td><td colspan="3">Plant structure, producer, chemical process, carbon, carbon dioxide, aerobic, anaerobic, molecule, sugars, photosynthesis</td></tr>
<tr><td colspan="4">Phenomena</td></tr>
<tr><td>Context/ Phenomena</td><td colspan="3">Some example phenomena for 5-LS1-1:
• A neoregelia plant sits on the branch of a much larger kapok tree in the Cloud Forest of South America.
• A plant grows in a classroom and the students weigh the soil every day. The weight of the soil does not change over time but the plant continues to grow.
• Spanish moss hangs from the branches of a live oak tree in the swamps of Louisiana.
• Strawberries sold in a supermarket were grown inside of a greenhouse without soil.</td></tr>
<tr><td colspan="4">This Performance Expectation and associated Evidence Statements support the following Task Demands.</td></tr>
<tr><td colspan="4">Task Demands</td></tr>
<tr><td colspan="4">1. Sort observations into those that appear to support competing (given) arguments, or into those that support, contradict, or are not relevant to a given argument. Observations are from animations, simulations, or other given material.</td></tr>
<tr><td colspan="4">2. Sort, tabulate, classify, separate, and/or categorize relevant from irrelevant evidence (observations) or data.</td></tr>
<tr><td colspan="4">3. Select from a given collection additional relevant observations that would help distinguish between competing arguments or the veracity of a single argument.</td></tr>
<tr><td colspan="4">4. Select, identify, or describe apparent counterexamples to a supported argument.</td></tr>
</table>

5. Identify from a given collection—or explain in writing—flaws in observation that lead to an apparent counterexample, or explain the counterexample in terms of grade-level appropriate properties of plant growth.
6. Sort statements into categories such as facts, reasonable judgments based on available facts, and speculation.
7. Articulate the evidence supporting and/or contradicting an argument that plants chiefly need air and water for growth.

<table>
<tr><td>Performance Expectation</td><td colspan="3">5-LS2-1
Develop a model to describe the movement of matter among plants, animals, decomposers, and the environment.</td></tr>
<tr><td>Dimensions</td><td>Developing and Using Models
• Develop a model to describe phenomena.</td><td>LS2.A: Interdependent Relationships in Ecosystems
• The food of almost any kind of animal can be traced back to plants. Organisms are related in food webs in which some animals eat plants for food and other animals eat the animals that eat plants. Some organisms, such as fungi and bacteria, break down dead organisms (both plants or plant parts and animals) and therefore operate as “decomposers.” Decomposition eventually restores (recycles) some materials back to the soil. Organisms can survive only in environments in which their particular needs are met. A healthy ecosystem is one in which multiple species of different types are each able to meet their needs in a relatively stable web of life. Newly introduced species can damage the balance of an ecosystem.

LS2.B: Cycles of Matter and Energy Transfer in Ecosystems
• Matter cycles between the air and soil and among plants, animals, and microbes as these organisms live and die. Organisms obtain gases and water from the environment and release waste matter (gas, liquid, or solid) back into the environment.</td><td>Systems and System Models
• A system can be described in terms of its components and their interactions.</td></tr>
<tr><td>Clarifications and Content Limits</td><td colspan="3">Clarification Statements
• Emphasis is on the idea that matter that is not food (air, water, decomposed materials in soil) is changed by plants into matter that is food.
• Examples of systems could include organisms, ecosystems, and Earth.

Content Limits
• <u>Assessment does not include</u>: molecular explanations.</td></tr>
<tr><td>Science Vocabulary Students Are Expected to Know</td><td colspan="3">Organism, bacteria, fungus, algae, gas, nutrients, producer, consumer, decomposer, cycle, conserve, products, relationship, waste, recycle, species, balance</td></tr>
<tr><td>Science Vocabulary Students Are Not Expected to Know</td><td colspan="3">Chemical process, reaction, molecule, carbon, carbon dioxide, oxygen, sugar, aerobic, anaerobic, photosynthesis</td></tr>
<tr><td colspan="4" align="center">Phenomena</td></tr>
<tr><td>Context/ Phenomena</td><td colspan="3">Some example phenomena for 5-LS2-1:
• Insects in a terrarium only survive when bacteria and plants are present.
• A new fish tank must rest for 2–3 weeks with water before introducing fish or the fish die.
• Under a microscope, a sample of soil contains many bacteria, but a sample of desert sand does not.
• Farmers put fish in stock tanks to keep them clean.</td></tr>
</table>

This Performance Expectation and associated Evidence Statements support the following Task Demands.	
Task Demands	
1. Select or identify from a collection of potential model components, including distractors, the parts of a model needed to describe the movement of matter among plants, animals, decomposers, and the environment.*	
2. Manipulate the components of a model to demonstrate properties, processes, and/or events that result in the movement of matter among plants, animals, decomposers, and the environment, including the relationships of organisms and/or the cycle(s) of matter and/or energy.	
3. Articulate, describe, illustrate, select, or identify the relationships among components of a model that describe the movement of matter among plants, animals, decomposers, and the environment.	
4. Make predictions about the effects of changes in model components, including the substitution, elimination, or addition of matter and/or an organism and the result.	

*denotes those task demands which are deemed appropriate for use in stand-alone item development

<table>
<tr><td>Performance Expectation</td><td colspan="3">5-ESS1-1
Support an argument that the apparent brightness of the sun and stars is due to their relative distances from Earth.</td></tr>
<tr><td>Dimensions</td><td>Engaging in Argument from Evidence
• Support an argument with evidence, data, or a model.</td><td>ESS1.A: The Universe and Its Stars
• The sun is a star that appears larger and brighter than other stars because it is closer. Stars range greatly in their distance from Earth.</td><td>Scale, Proportion, and Quantity
• Natural objects exist from the very small to the immensely big.</td></tr>
<tr><td>Clarifications and Content Limits</td><td colspan="3">Content Limits
• Assessment is limited to relative distances, not sizes, of stars.
• Assessment does not include other factors that affect apparent brightness (such as stellar masses, age, stage, etc.).
• Assessment does not include absolute brightness.
• <u>Students do not need to know:</u>
o Specific stars and their names.
o Luminosity and how that is affected by the size/age of a star.
o Flux or how to calculate it.</td></tr>
<tr><td>Science Vocabulary Students Are Expected to Know</td><td colspan="3">Space, planet, sun’s size, solar system, moon, burn, star brightness, constellation, galaxy, visible, astronomical.</td></tr>
<tr><td>Science Vocabulary Students Are Not Expected to Know</td><td colspan="3">Lunar phase, eclipse, celestial, mass, comet, light year, astronomical unit, emit, interstellar, fission, fusion, radiation, spectrum, star size, star composition, star formation, star types, luminosity, flux.</td></tr>
<tr><td colspan="4">Phenomena</td></tr>
<tr><td>Context/ Phenomena</td><td colspan="3">Some example phenomena for 5-ESS1-1:
• Most stars cannot be seen during the daytime but can be seen at night.
• The sun is never seen at the same time as other stars in the sky.
• Alpha Centauri A is larger than the sun but does not look as bright in the sky.
• Street lights that are farther away from you look dimmer.</td></tr>
<tr><td colspan="4">This Performance Expectation and associated Evidence Statements support the following Task Demands.</td></tr>
<tr><td colspan="4">Task Demands</td></tr>
<tr><td colspan="4">1. Organize, arrange (e.g., using illustrations and/or labels), or summarize data to highlight trends, patterns, or correlations in how the brightness of stars is based on their relative distance from Earth.*</td></tr>
<tr><td colspan="4">2. Generate/construct graphs, tables, or assemblages of illustrations and/or labels of data that document patterns, trends, or correlations in how the brightness of stars is based on their relative distance from Earth. This may include sorting out distractors.*</td></tr>
<tr><td colspan="4">3. Describe, identify, and/or select information needed to support an explanation.*</td></tr>
<tr><td colspan="4">4. Use relationships identified in the data to predict the distance of a star depending on its brightness, or vice versa.*</td></tr>
</table>

5. Identify patterns or evidence in the data that supports inferences about how the brightness of stars depends on their relative distance from Earth.

*denotes those task demands which are deemed appropriate for use in stand-alone item development

This page intentionally left blank.

Performance Expectation	**5-ESS1-2** Represent data in graphical displays to reveal patterns of daily changes in length and direction of shadows, day and night, and the seasonal appearance of some stars in the night sky.		
Dimensions	**Analyzing and Interpreting Data** • Represent data in graphical displays (bar graphs, pictographs, and/or pie charts) to reveal patterns that indicate relationships.	**ESS1.B: Earth and the Solar System** • The orbits of Earth around the sun and of the moon around Earth, together with the rotation of Earth about an axis between its North and South poles, cause observable patterns. These include day and night; daily changes in the length and direction of shadows; and different positions of the sun, moon, and stars at different times of the day, month, and year.	**Patterns** • Similarities and differences in patterns can be used to sort, classify, communicate, and analyze simple rates of change for natural phenomena.
Clarifications and Content Limits	**Content Limits** • Examples of patterns could include the position and motion of Earth with respect to the sun and selected stars that are visible only in particular months. • While the names of celestial objects, stars, or constellations can be included, students are not expected to identify them. • Objects to be used to assess this PE are limited to the sun, Earth’s moon, Earth, and stars/constellations visible in Earth’s night sky. • “Positions of the moon” refers to its location in Earth’s sky and not its appearance (phase). • Assessment does not include cause of seasons, lunar phases, or the position of the sun in the sky throughout the year.		
Science Vocabulary Students Are Expected to Know	Circular motion, universe, Earth’s rotation, galaxy, axis, solar system, Milky Way, constellation, moon phases, lunar astronomical, orbit, tilt, annual, rotation, revolution.		
Science Vocabulary Students Are Not Expected to Know	Eclipse, celestial, comet, light year, astronomical unit, stellar.		
Phenomena			
Context/ Phenomena	Some example phenomena for 5-ESS1-2: • The shadow cast by a sundial changes position and size throughout the day. • A constellation that is viewed right above someone’s house at 8:00 p.m. one night can no longer be seen at 8:00 p.m. in a few months. • The sun is seen in the sky only during the day • It gets dark out after the sun goes below the horizon.		

This Performance Expectation and associated Evidence Statements support the following Task Demands.
Task Demands
1. Organize, arrange (e.g., using illustrations and/or labels), or summarize data to highlight trends, patterns, or correlations in how the data changes over time.*
2. Generate/construct graphs, tables, or groups of illustrations and/or labels of data that document patterns, trends, or correlations in how the data change over time. This may include sorting out distractors.*

3. Use relationships identified in the data to predict whether or not the pattern will continue OR how the data will look at some time in the future.*
4. Identify patterns or evidence in the data that supports inferences about the phenomena.

*denotes those task demands which are deemed appropriate for use in stand-alone item development

<table>
<tr><td>Performance Expectation</td><td colspan="3">5-ESS2-1
Develop a model using an example to describe ways the geosphere, biosphere, hydrosphere, and/or atmosphere interact.</td></tr>
<tr><td>Dimensions</td><td>Developing and Using Models
• Develop a model using an example to describe a scientific principle.</td><td>ESS2.A: Earth Materials and Systems
• Earth's major systems are the geosphere (solid and molten rock, soil, and sediments), the hydrosphere (water and ice), the atmosphere (air), and the biosphere (living things, including humans). These systems interact in multiple ways to affect Earth's surface materials and processes. The ocean supports a variety of ecosystems and organisms, shapes landforms, and influences climate. Winds and clouds in the atmosphere interact with the landforms to determine patterns of weather.</td><td>Systems and System Models
• A system can be described in terms of its components and their interactions.</td></tr>
<tr><td>Clarifications and Content Limits</td><td colspan="3">Clarification Statements
• Examples could include the influence of the ocean on ecosystems, landform shape, and climate; the influence of the atmosphere on landforms and ecosystems through weather and climate; and the influence of mountain ranges on winds and clouds in the atmosphere.
• The geosphere, hydrosphere, atmosphere, and biosphere are each a system.

Content Limits
• Assessment is limited to the interactions of two systems at a time.</td></tr>
<tr><td>Science Vocabulary Students are Expected to Know</td><td colspan="3">core, mantle, crust, solid, liquid, gas, vapor, tundra, boreal forest, deciduous forest, grassland, desert, savannah, tropical rainforest, freshwater, marine, high pressure, low pressure, currents, circulation</td></tr>
<tr><td>Science Vocabulary Students are Not Expected to Know</td><td colspan="3">troposphere, stratosphere, mesosphere, thermosphere, ionosphere, chaparral</td></tr>
<tr><td colspan="4">Phenomena</td></tr>
<tr><td>Context/ Phenomena</td><td colspan="3">Some example phenomena for 5-ESS2-1:
• The land area found on the beaches around Nantucket Sound in 2016 were about three times the land area in the same location in 1984.
• In 2016, Tucson, Arizona received more rain between June and September than Yuma, Arizona received during the entire year.
• The amount of carbon dioxide in the atmosphere measured at Mauna Loa Observatory in April is 397 parts per million. The amount measured at the same location the previous September was 2% less.
• In 1980, the salt content in the freshwater Biscayne Aquifer in Florida was 50 milligrams per liter. In 1997, the salt content of the same water was 1,000 milligrams per liter.</td></tr>
<tr><td colspan="4">This Performance Expectation and associated Evidence Statements support the following Task Demands.</td></tr>
<tr><td colspan="4">Task Demands</td></tr>
<tr><td colspan="4">1. Select or identify from a collection of potential model components, including distractors, the components needed to model the phenomenon. Components might include labels, text, steps in a process.</td></tr>
</table>

2. Assemble or complete, from a collection of potential model components, an illustration or flow chart that is capable of representing how the geosphere, biosphere, hydrosphere, and/or atmosphere interact. This <u>does not</u> include labeling an existing diagram. 3. Manipulate the components of a model to demonstrate the changes, properties, processes, and/or events that act to result in the phenomenon.
4. Make predictions about the effects of changes in the geosphere, biosphere, hydrosphere, or atmosphere on each other. Predictions can be made by manipulating model components, completing illustrations, or selecting from lists with distractors.
5. Given models or diagrams of ways in which the geosphere, biosphere, hydrosphere, and/or atmosphere interact, identify relationships between the spheres and how a change in one causes a change in another.
6. Identify missing components, relationships, or other limitations of the model.

<table>
<tr><td>Performance Expectation</td><td colspan="3">5-ESS2-2
Describe and graph the amounts of salt water and fresh water in various reservoirs to provide evidence about the distribution of water on Earth.</td></tr>
<tr><td>Dimensions</td><td>Using Mathematics and Computational Thinking
• Describe and graph quantities such as area and volume to address scientific questions.</td><td>ESS2.C: The Roles of Water in Earth’s Surface Processes
• Nearly all of Earth’s available water is in the ocean. Most fresh water is in glaciers or underground; only a tiny fraction is in streams, lakes, wetlands, and the atmosphere.</td><td>Scale, Proportion, and Quantity
• Standard units are used to measure and describe physical quantities such as weight and volume.</td></tr>
<tr><td>Clarifications and Content Limits</td><td colspan="3">Content Limits
• Assessment is limited to oceans, lakes, rivers, glaciers, ground water, and polar ice caps, and does not include the atmosphere.
• Students will not be provided a calculator.</td></tr>
<tr><td>Science Vocabulary Students are Expected to Know</td><td colspan="3">Cycle, fresh water, glacial movement, global, ground water, moisture, polar ice caps, properties of soil, reservoir, soil composition, water capacity, feature, glacial, hydrosphere, surface feature, water cycle, wetland.</td></tr>
<tr><td>Science Vocabulary Students are Not Expected to Know</td><td colspan="3">Coastal, crust, internal, distribution, hydrological cycle, percentage</td></tr>
<tr><td colspan="4" align="center">Phenomena</td></tr>
<tr><td>Context/ Phenomena</td><td colspan="3">The phenomenon for these PEs are the given data. Samples of phenomena should describe the data set(s) to be given in terms of patterns or relationships to be found in the data, and the columns and rows of a hypothetical table presenting the data, even if the presentation is not tabular. The description of the phenomenon should describe the presentation format of the data (e.g., maps, tables, graphs, etc.). For this performance expectation the phenomena are a set of data on the relative volume of water in different reservoirs on Earth using standard units for weight or volume.

Some example sets of data for 5-ESS2-2:
• Melting ice from the Arctic ice cap is currently adding fresh water to the very salty Arctic Ocean.
• Melting ice from the Greenland Ice Sheet is currently adding fresh water to the very salty Arctic Ocean.
• The Potomac River in the eastern United States is tidally influenced over XX% of its length. This tidal influence from the ocean results in the portion of the river near the ocean being a mixture of salt and fresh water and the portion of the river far from the ocean being fresh water.
• Salt water intrusion on Cape Cod, Florida, or California.</td></tr>
<tr><td colspan="4" align="center">This Performance Expectation and associated Evidence Statements support the following Task Demands.</td></tr>
<tr><td colspan="4" align="center">Task Demands</td></tr>
<tr><td colspan="4">1. Illustrate, graph, or identify relevant features or data that can be used to calculate or estimate relationships between the relative volumes of water in different reservoirs on Earth.</td></tr>
</table>

2. Calculate or estimate properties or relationships of the relative volumes of water in different reservoirs on Earth, based on data from one or more sources.
3. Compile, from given information, the data needed for a particular inference about the relative volumes of water in different reservoirs on Earth. This can include sorting out the relevant data from the given information.

<table>
<tr><td>Performance Expectation</td><td colspan="3">5-ESS3-1
Obtain and combine information about ways individual communities use science ideas to protect the Earth’s resources and environment.</td></tr>
<tr><td>Dimensions</td><td>Obtaining, Evaluating, and Communicating Information
• Obtain and combine information from books and/or other reliable media to explain phenomena or solutions to a design problem.</td><td>ESS3.C: Human Impacts on Earth Systems
• Human activities in agriculture, industry, and everyday life have had major effects on the land, vegetation, streams, ocean, air, and even outer space. But individuals and communities are doing things to help protect Earth’s resources and environments.</td><td>Systems and System Models
• A system can be described in terms of its components and their interactions.</td></tr>
<tr><td>Clarifications and Content Limits</td><td colspan="3"></td></tr>
<tr><td>Science Vocabulary Students are Expected to Know</td><td colspan="3">Atmosphere, cycle, fresh water, global, ground water, moisture, polar ice caps, properties of soil, soil composition, water cycle</td></tr>
<tr><td>Science Vocabulary Students are Not Expected to Know</td><td colspan="3">Coastal, crust, internal, distribution, hydrological cycle, reservoir, glacial movement, water capacity, glacial, hydrosphere, reservoir, feature, surface feature, wetland, percentage</td></tr>
<tr><td colspan="4">Phenomena</td></tr>
<tr><td>Context/ Phenomena</td><td colspan="3">Engineering practices are built around meaningful design problems rather than phenomena. For this PE, there are 2 phenomena and 2 design problems.

Some example phenomena for 5-ESS3-1:
• In England in 1965, there were about 182,000 bee colonies. By 2010, there were about 83,000 bee colonies.
• There is a haze in the air in Beijing, China’s capital city, which makes it hard to see long distances. The haze becomes worse on cold winter days.

Some example design problems for 5-ESS3-1:
• A company is going to put a new logging road in an area where grizzly bears live. The US Forest Service tells them that they need to pay attention to where they are going to put the road. The path of the road should be chosen so that it doesn’t disturb grizzly bear habitat very much.
• A flower garden to attract honeybees is being built. The type and color of flowers, garden placement, flower placement, and other features are chosen to attract honeybees.</td></tr>
<tr><td colspan="4">This Performance Expectation and associated Evidence Statements support the following Task Demands.</td></tr>
<tr><td colspan="4">Task Demands</td></tr>
<tr><td colspan="4">1. Identify, evaluate, combine, organize, and/or communicate information (from texts, illustrations, animations, simulations, tables, or graphs) that is needed to make an informed decision related to human impacts on natural systems, solve a particular design problem, or complete a specified task.</td></tr>
</table>

2. Assemble or complete an illustration, graph, set of labels, or a flow chart that shows how the various pieces of information, which are needed to make an informed decision, solve a particular design problem, or complete a specified task, are interrelated. This does not include labeling an existing diagram.
3. Identify patterns or evidence in the data that supports inferences about human impacts on natural systems or a particular solution to a design problem or task.
4. Examine, identify or select positive or negative effects/implications of a community idea or design problem. This would include identifying potential positive or negative effects, especially when dealing with design solutions, and classifying the effects/implications as positive or negative and supporting those classifications with the relevant data.
5. Formulate a design or make an inference or conclusion, based on identified or combined information, evidence or data related to human impacts on natural systems, solution of a particular design problem, or completion of a specified task.
6. Evaluate a design or make an inference or conclusion, based on identified or combined information, evidence or data related to human impacts on natural systems, solution of a particular design problem, or completion of a specified task.

Specifications for Middle School

Introduction

This document presents *cluster specifications* for use with the Next Generation Science Standards (NGSS). These standards are based on the Framework for K-12 Science Education. The present document is not intended to replace the standards, but rather to present guidelines for the development of items and item clusters used to measure those standards.

The remainder of this section provides a very brief introduction to the standards and the framework, an overview of the design and intent of the item clusters, and a description of the cluster specifications that follow. The bulk of the document is composed of cluster specifications, organized by grade and standard.

Background on the framework and standards

The Framework for K-12 Science Education are organized around three core dimensions of scientific understanding. The standards are derived from these same dimensions:

- Disciplinary Core Ideas: The fundamental ideas that are necessary for understanding a given science discipline. The core ideas all have broad importance within or across science or engineering disciplines, provide a key tool for understanding or investigating complex ideas and solving problems, relate to societal or personal concerns, and can be taught over multiple grade levels at progressive levels of depth and complexity.
- Science and Engineering Practices: The practices are what students DO to make sense of phenomena. They are both a set of skills and a set of knowledge to be internalized. The SEPs reflect the major practices that scientists and engineers use to investigate the world and design and build systems.
- Cross-Cutting Concepts: These are concepts that hold true across the natural and engineered world. Students can use them to make connections across seemingly disparate disciplines or situations, connect new learning to prior experiences, and more deeply engage with material across the other dimensions. The NGSS requires that students explicitly use their understanding of the CCCs to make sense of phenomena or solve problems.
- There is substantial overlap between and among the three dimensions. For example, the cross-cutting concepts are echoed in many of the disciplinary core ideas. The core ideas are often closely intertwined with the practices. This overlap reflects the nature of science itself. For example, we often come to understand and communicate causal relationships by employing models to make sense of observations. Even within a dimension, overlap exists. Quantifying characteristics of phenomena is important in developing an understanding of them, so employing computational and mathematical thinking in the construction and use of models is a very common scientific practice, and one of the cross-cutting concepts suggests that scientists often infer causality by observing patterns. In short, the dimensions are not orthogonal.

The framework envisions effective science education as occurring at the intersection of these interwoven dimensions: students learn science by doing science—applying the practices through the lens of the cross-cutting concepts to investigate phenomena that relate to the content of the disciplinary core ideas.

Item clusters

Each item cluster is designed to engage the examinee in a grade-appropriate, meaningful scientific activity aligned to a specific standard.

Each cluster begins with a phenomenon, an observable fact or design problem that engages student interest and can be explained, modeled, investigated, or designed using the knowledge and skill described by the standard in question.

What it means to be observable varies across practices. For example, a phenomenon for a performance expectation exercising the a*nalyze data* practice may be observable through regularities in a data set, while standards related to the *development and use of models* might be something that can be watched, seen, felt, smelled, or heard.

What it means to be observable also varies across grade levels. For example, elementary-level phenomena are very concrete and directly observable. At the high school level, an observation of the natural world may be more abstract--for example, "observing" changes in the chemical composition of cells through the observation of macroscopic results of those changes on organism physiology, or through the measurement of system- or organ-level indications.
Content limits refine the intent of the performance expectations and provide limits on what may be asked of items in the cluster to structure the student activity. The content limits also reflect the disciplinary core ideas learning progressions that are present in the K-12 Framework for Science Education.

The task or goal should be explicitly stated in the stimulus or the first item in the cluster: statements such as "In the questions that follow, you will develop a model that will allow you to identify moons of Jupiter," or "In the questions below, you will complete a model to describe the processes that lead to the steam coming out of the teapot."
Whereas item clusters have been described elsewhere as "scaffolded," they are better described as providing structure to the task. For example, some clusters begin with students summarizing data to discover patterns that may have explanatory value. Depending on the grade level and nature of the standard, items may provide complete table shells or labeled graphs to be drawn, or may require the student to choose what to tabulate or graph. Subsequent items may ask the student to note patterns in the tabulated or graphed data and draw on domain content knowledge to posit explanations for the patterns.
These guidelines for clusters do not appear separately in the specifications. Rather, they apply to all clusters.

Structure of the cluster specifications

The item cluster specifications are designed to guide the work of item writers and the review of item clusters by stakeholders.
Each item cluster has the following elements:

- The text of the performance expectations, including the practice, core idea, and cross-cutting concept.
- Content limits, which refine the intent of the performance expectations and provide limits of what may be asked of examinees. For example, they may identify the specific formulae that students are expected to know or not know.
- Vocabulary, which identifies the relevant technical words that students are expected to know, and related words that they are explicitly not expected to know. The vocabulary included in both sections (Students are Expected to Know and Students Are Not Expected to Know) were developed after the reviews of standards at the current/preceding grades, the original NGSS documentation, and item writer reference documentation including the Children Writers' Word Book and ED Core Vocabularies in Reading, Mathematics, Science and Social Studies. All vocabulary included in the specifications was reviewed and edited by teacher committees during the specification reviews by states. Of course, the latter category should not be considered exhaustive, since the boundaries of relevance are ambiguous, and the list is limited by the imagination of the writers/reviewers.
- Sample phenomena, which provide some examples of the sort of phenomena that would support effective item clusters related to the standard in question. In general, these should be guideposts, and item writers should seek comparable phenomena, rather than drawing on those within the documents. Novelty is valued when applying scientific practices.
- Task demands comprise the heart of the specifications. These statements identify the types of items and activities that item writers should use, and each item written should be clearly linked to one or more of the demands. The verbs in the demands (e.g., *select, identify, illustrate, describe*) provide guidance on the types of interactions that item writers might employ to elicit the student response. We avoid explicitly identifying interaction types or item formats to accommodate future innovations and to avoid discouraging imaginative work by the item writers. All task demands should be aligned to a minimum of one of the three dimensions (DCI, SEP and CCCs) and across task demands within a cluster, all three dimensions must be addressed.

- For each cluster we present, the printed documentation includes the cluster, the task demands represented by each item, and its linkage to the content, practice and cross-cutting concept identified in the performance expectation.

Item cluster specifications follow, organized by domain and standard.

<table>
<tr><td>Performance Expectation</td><td colspan="3">MS-PS1-1
Develop models to describe the atomic composition of simple molecules and extended structures.</td></tr>
<tr><td>Dimensions</td><td>Developing and Using Models
• Develop and/or use a model to predict and/or describe phenomena.</td><td>PS1.A: Structure and Properties of Matter
• Substances are made from different types of atoms, which combine with one another in various ways. Atoms form molecules that range in size from two to thousands of atoms.
• Solids may be formed from molecules, or they may be extended structures with repeating subunits (e.g., crystals).</td><td>Scale, Proportion, and Quantity
• Time, space, and energy phenomena can be observed at various scales, using models to study systems that are too large or too small.</td></tr>
<tr><td>Clarifications and Content Limits</td><td colspan="3">Clarification Statements
• Emphasis is on identifying elements vs. compounds and their basic units of atoms and molecules.
• Emphasis is on developing models of molecules that vary in complexity.
• Examples of simple molecules could include ammonia, methanol, methane, water, carbon dioxide, etc.
• Examples of molecular-level models could include drawings, 3D ball and stick structures, or computer representations showing different molecules with different types of atoms.
• Examples of extended structures could include sodium chloride or diamonds.

Content Limits
• Assessment does not include valence electrons and bonding energy, discussing the ionic nature of subunits of complex structures, or a complete description of all individual atoms in a complex molecule or extended structure is not required.
• Modelling should be limited to molecules that have only one type of bond, no combination of bonds; the structure of the molecule is easy to model; single bonded molecules.
• Students are not expected to memorize the atomic characteristics of any element.
• <u>Students do not need to know</u>: valence electrons and bonding energy, discussing the ionic nature of subunits of complex structures, a complete description of all individual atoms in a complex molecule or extended structure, memorization of atoms found in different molecule, VSEPR or geometric arrangements, the difference between single, double, and triple bonding, periodic table patterns and how it affects bonding, oxidation numbers, polyatomic ions.</td></tr>
<tr><td>Science Vocabulary Students are Expected to Know</td><td colspan="3">Element, compound, mixtures, homogenous, heterogeneous, pure substances, solution, solvent, solute.</td></tr>
<tr><td>Science Vocabulary Students are Not Expected to Know</td><td colspan="3">Valence electrons, subatomic particles such as protons, electrons, neutrons, neutrinos etc., ions, positive or negative charges, covalent bond, ionic bond.</td></tr>
<tr><td colspan="4" align="center">Phenomena</td></tr>
<tr><td>Context/ Phenomena</td><td colspan="3">Some example phenomena for MS-PS1-1:</td></tr>
</table>

	• Submarines can stay underwater for months using sea water as a source of oxygen for air. Special machines run electricity through large amounts of sea water, generating oxygen from the water. • Water and hydrogen peroxide are both made up of hydrogen and oxygen. When water is poured on a chunk of $CaCO_3$, there is no reaction. When hydrogen peroxide is poured on a chunk of $CaCO_3$, it fizzes. • Oxygen (O_2) is a gas we breathe to stay alive. Ozone (O_3), also made only of oxygen atoms, is unhealthy to breathe.

This Performance Expectation and associated Evidence Statements support the following Task Demands.
Task Demands
1. Identify or assemble from a collection of potential model components, including distractors, components of a model that describes the structures of atoms, molecules, or extended molecules and/or how they interact, or explains how atoms of the same or different element(s) are arranged in repeated patterns in extended structures.
2. Describe, select, and/or identify the relationships among components of a model that describes the structures of atoms, molecules, or extended molecules and/or how they interact, or explains how atoms of the same or different element(s) are arranged in repeated patterns in extended structures.
3. Assemble, illustrate, describe, and/or complete a model or manipulate components of a model to describe the structure of an atom, molecule, or extended molecule and/or how they interact, or to explain or predict how atoms of the same or different element(s) are arranged in repeated patterns in extended structures.

<table>
<tr><td>Performance Expectation</td><td colspan="3">MS-PS1-2
Analyze and interpret data on the properties of substances before and after the substances interact to determine whether a chemical reaction has occurred.</td></tr>
<tr><td>Dimensions</td><td>Analyzing and Interpreting Data
• Analyze and interpret data to determine similarities and differences in findings.</td><td>PS1.A: Structure and Properties of Matter
• Each pure substance has characteristic physical and chemical properties (for any bulk quantity under given conditions) that can be used to identify it.

PS1.B: Chemical Reactions
• Substances react chemically in characteristic ways. In a chemical process, the atoms that make up the original substances are regrouped into different molecules, and these new substances have different properties from those of the reactants.</td><td>Patterns
• Macroscopic patterns are related to the nature of microscopic and atomic-level structure.</td></tr>
<tr><td>Clarifications and Content Limits</td><td colspan="3">Clarification Statements
• Examples of reactions could include burning sugar or steel wool, fat reacting with sodium hydroxide, and mixing zinc with hydrogen chloride.

Content Limits
• Assessment is limited to analysis of the following properties: density, melting point, boiling point, solubility, flammability, and odor.
• Students are not expected to balance chemical equations or to determine whether a chemical equation is balanced or not.
• Students are expected to know that mass/matter is neither destroyed nor created.</td></tr>
<tr><td>Science Vocabulary Students are Expected to Know</td><td colspan="3">Reactant, product, compound, matter, mass, volume, density, melting point, boiling point, freezing point, solubility, dissolve, flammability, odor, gas, solid, liquid, chemical bonds.</td></tr>
<tr><td>Science Vocabulary Students are Not Expected to Know</td><td colspan="3">Collision theory, oxidation, reduction, intramolecular attractions, intermolecular attractions, solvent, solute, precipitant, limiting reactant, excess reactant, covalent bond, ionic bond, rate of reaction, acid, base, salt (as an ionic crystal), fusion, fission, homogeneous mixture, heterogeneous mixture.</td></tr>
<tr><td colspan="4" align="center">Phenomena</td></tr>
<tr><td>Context/ Phenomena</td><td colspan="3">For this performance expectation the phenomena are mixtures of substances that provide sets of data. Those are the observations and/or measurements concerning the physical and chemical properties of the involved substances before and after mixing that the kids will look at to discover patterns. Below, we enumerate some of the mixtures that might provide the data sets to be analyzed.

All phenomenon for this PE should be situations where a chemical reaction is not immediately apparent.

Some example phenomena for MS-PS1-2:
• Rainwater can produce stains on car paint. Reports of these stains are more common in the Southeastern U.S. than they are in the Midwest.</td></tr>
</table>

	• Portions of marble statues that are exposed to rainwater crack and crumble over time. Portions of marble statues that are sheltered develop a black coating over time. • When sugar crystals are added to vinegar in a bowl, the crystals disappear. When crystals of table salt are added to vinegar in a bowl, the mixture begins to bubble and foam. • Table sugar exposed to an open flame transforms into a gooey, dark substance. Wood exposed to an open flame transforms into ash.
This Performance Expectation and associated Evidence Statements support the following Task Demands.	
Task Demands	
1. Organize, arrange, and/or generate/construct graphs, tables, or assemblages of illustrations and/or labels of data that document patterns, trends, or correlations among observations and data concerning the physical and chemical properties of the substances involved. This may include sorting out distractors.	
2. Describe and/or summarize data (e.g., using illustrations and/or labels), to identify/highlight trends, patterns, or correlations among observations and data concerning the physical and chemical properties of the beginning and ending substances being investigated.*	
3. Use relationships identified in the data to predict whether the mixing of substances similar to the ones under study will result in the occurrence of a chemical reaction.	
4. Identify patterns or evidence in the data that support inferences about any changes that occurred in the microscopic or atomic-level arrangements of the substances involved.*	

*denotes those task demands which are deemed appropriate for use in stand-alone item development

Performance Expectation	**MS-PS1-3** Gather and make sense of information to describe that synthetic materials come from natural resources and impact society.		
Dimensions	**Obtaining, Evaluating, and Communicating Information** • Gather, read, and synthesize information from multiple appropriate sources and assess the credibility, accuracy, and possible bias of each publication and methods used, and describe how they are supported or not supported by evidence.	**PS1.A: Structure and Properties of Matter** • Each pure substance has characteristic physical and chemical properties (for any bulk quantity under given conditions) that can be used to identify it. **PS1.B: Chemical Reactions** • Substances react chemically in characteristic ways. In a chemical process, the atoms that make up the original substances are regrouped into different molecules, and these new substances have different properties from those of the reactants.	**Structure and Function** • Structures can be designed to serve particular functions by taking into account properties of different materials, and how materials can be shaped and used.
Clarifications and Content Limits	**Clarification Statements** • Emphasis is on natural resources that undergo a chemical process to form the synthetic material. • Examples of new materials could include new medicine, foods, building materials, plastics and alternative fuels **Content Limits** • Assessment is limited to qualitative information. • Students are not required to know particular names for synthetic materials (i.e. rayon, polyester, acrylic, nylon, rayon, acetate, orlon, Kevlar) • <u>Students do not need to know</u>: the types of reaction mechanisms involved in chemical reactions such as polymerization.		
Science Vocabulary Students are Expected to Know	Atom, molecule, pure substance, subunit, molecular arrangement, matter, particle, pressure, conductivity, reactant, dissolve, mineral, conductive, separation method (for mixtures), sodium chloride, carbon dioxide, negative impact, petroleum, natural gas, oil		
Science Vocabulary Students are Not Expected to Know	Acid, base, reversible reactions, irreversible reactions, condensation reaction, polymer, polymerization, bond, electron configuration, chromatography, catalyst, electron transfer, graphite, pharmaceutical, synthetic polymer, harvesting of resources, oil shale, geopolitical, extract, cost-benefit, organic materials		
Phenomena			
Context/ Phenomena	Some example phenomena for MS-PS1-3: • It is difficult for the naked eye to tell the difference between cubic zirconia (CZ) and diamond, but a genuine diamond will give off a strong blue fluorescence when held under U.V. light. • Naturally occurring penicillin from penicillium mold is an effective antibiotic against infections, but it is broken up by stomach acid and can only be injected into the bloodstream. • The bark of the white willow tree can be used as an alternative to aspirin for pain relief.		

	• Nylon and Kevlar are both synthetic fabrics, but Kevlar is much stronger – about five times as strong as steel.

This Performance Expectation and associated Evidence Statements support the following Task Demands.
Task Demands
1. Analyze and interpret scientific evidence from multiple scientific/technical sources including text, diagrams, charts, symbols and mathematical representations to describe how synthetic materials are made and how they come from natural resources.*
2. Based on the information provided, identify, describe or illustrate a claim regarding the relationship between a characteristic of a synthetic material and its function in real world applications.
3. Identify, summarize, or organize given data or other information to support or refute a claim that relates characteristic of a synthetic material to its function in real world.
4. Identify relationships or patterns in scientific evidence at macroscopic and/or microscopic scales.
5. Synthesize an explanation that incorporates the scientific evidence from multiple sources.
6. Using scientific evidence, evaluate the validity/relevance/reliability of using synthetic materials as alternatives to natural materials and/or their impact on society.

*denotes those task demands which are deemed appropriate for use in stand-alone item development
**For stand-alone items, focus on charts, diagrams, etc. rather than text-heavy stems for time considerations.

<table>
<tr><td>Performance Expectation</td><td colspan="3">MS-PS1-4
Develop a model that predicts and describes changes in particle motion, temperature, and state of a pure substance when thermal energy is added or removed.</td></tr>
<tr><td>Dimensions</td><td>Developing and Using Models
• Develop a model to predict and/or describe phenomena.</td><td>PS1.A: Structure and Properties of Matter
• Gases and liquids are made of molecules or inert atoms that are moving about relative to each other.
• In a liquid, the molecules are constantly in contact with others; in a gas, they are widely spaced except when they happen to collide. In a solid, atoms are closely spaced and may vibrate in position but do not change relative locations.
• The changes of state that occur with variations in temperature or pressure can be described and predicted using these models of matter.

PS3.A: Definitions of Energy
• The term “heat” as used in everyday language refers both to thermal energy (the motion of atoms or molecules within a substance) and the transfer of that thermal energy from one object to another. In science, heat is used only for this second meaning; it refers to the energy transferred due to the temperature difference between two objects. (secondary)
• The temperature of a system is proportional to the average internal kinetic energy and potential energy per atom or molecule (whichever is the appropriate building block for the system’s material). The details of that relationship depend on the type of atom or molecule and the interactions among the atoms in the material. Temperature is not a direct measure of a system's total thermal energy. The total thermal energy (sometimes called the total internal energy) of a system depends jointly on the temperature, the total number of atoms in the system, and the state of the material (secondary).</td><td>Cause and Effect
• Cause and effect relationships may be used to predict phenomena in natural or designed systems.</td></tr>
<tr><td>Clarifications and Content Limits</td><td colspan="3">Clarification Statements
• Emphasis is on qualitative molecular-level models of solids, liquids, and gases to show that adding or removing thermal energy increases or decreases kinetic energy of the particles until a change of state occurs.
• Examples of models could include drawings and diagrams.
• Examples of particles could include molecules of inert atoms.
• Examples of pure substances could include water, carbon dioxide, and helium.

Content Limits:
• Physical changes should be limited to freezing, melting, condensation, and evaporation.
• Assessment does not include:
o Sublimation (solid change of state directly to a gas);
o Calculations for internal energy, transfer of heat (q), (system and surroundings), entropy, work, and Hess’s law;
o Ideal gas laws and their relationships (Boyle’s, Charles, Combined, PV=nRT, etc.);
o The role that pressure and force (N) have in the kinetic molecular theory;
o Energy needed to break and reform chemical bonds in a chemical reaction, including the use of a catalyst to speed up a reaction;
o Absolute zero and kelvin (Celsius and Fahrenheit temperature only).</td></tr>
</table>

<table>
<tr><td></td><td>
<u>Students do not need to know</u>:

Atomic structure (electrons orbit around a nucleus containing protons and neutrons)
The structure and interactions of matter at the bulk scale are determined by electrical forces within and between atoms.
Stable forms of matter are those in which the electric and magnetic field energy is minimized.
A stable molecule has less energy, by an amount known as the binding energy, than the same set of atoms separated; one must provide at least this energy in order to break the bonds of a molecule.
That there is a single quantity called energy is due to the fact that a system’s total energy is conserved, even as, within the system, energy is continually transferred from one object to another and among its various possible forms.

</td></tr>
<tr><td>Science Vocabulary Students are Expected to Know</td><td>Phase, phase change, thermal energy, kinetic energy, pure substance, compound, thermometer, matter, melting, freezing, condensation, vapor, heat, vibrate, collide, inert atom.</td></tr>
<tr><td>Science Vocabulary Students are Not Expected to Know</td><td>Entropy, enthalpy, ideal gas law, sublimation, plasma, triple point, critical point, proton, neutron, electron, valence electrons, electrical energy, bond energy.</td></tr>
<tr><td colspan="2">Phenomena</td></tr>
<tr><td>Context/ Phenomena</td><td>Some example phenomena for MS-PS1-4:

A tea kettle is sitting on a stove, under heat. As the water in the kettle begins to boil, a stream of steam is visible outside of its spout.
Dew forms on the grass in the morning.
As sugar is heated in a pan, it turns from a white solid to a light brown liquid.
</td></tr>
<tr><td colspan="2">This Performance Expectation and associated Evidence Statements support the following Task Demands.</td></tr>
<tr><td colspan="2">Task Demands</td></tr>
<tr><td colspan="2">1. Select or identify from a collection of potential model components, including distractors, the components needed to model of the model changes in particle motion, temperature, and state of a pure substance when thermal energy is added or removed. Components might include: energy source, particles in motion, and boundaries of system.</td></tr>
<tr><td colspan="2">2. Assemble or complete, from a collection of potential model components, an illustration or flow chart that is capable of representing changes in particle motion, temperature, and state of a pure substance when thermal energy is added or removed. This <u>does not</u> include labeling an existing diagram.</td></tr>
<tr><td colspan="2">3. Manipulate the components of the model to demonstrate the changes, properties, processes, and/or events that act to result in the changes in particle motion, temperature, and state of a pure substance when thermal energy is added or removed.*</td></tr>
<tr><td colspan="2">4. Make predictions about the effects of changes in particle motion, temperature, and state of a pure substance when thermal energy is added or removed. Predictions can be made by manipulating model components, completing illustrations, or selecting from lists with distractors.*</td></tr>
<tr><td colspan="2">5. Given models or diagrams of particle motion, temperature, and state of a pure substance when thermal energy is added or removed, identify how they change over time in a given scenario OR identify the properties of the variables that cause the changes.</td></tr>
</table>

6. Identify missing components, relationships or other limitations of the model.
7. Describe, select, or identify the relationships among components of a model that describe changes in particle motion, temperature, and state of a pure substance when thermal energy is added or removed.

*denotes those task demands which are deemed appropriate for use in stand-alone item development
**TD4 must be used with TD3 (…by completing illustrations…etc. is what makes this need to be paired)

Performance Expectation	**MS-PS1-5** Develop and use a model to describe how the total number of atoms does not change in a chemical reaction and thus mass is conserved.		
Dimensions	**Developing and Using Models** • Develop and use a model to describe unobservable mechanisms.	**PS1.B: Chemical Reactions** • Substances react chemically in characteristic ways. In a chemical process, the atoms that make up the original substances are regrouped into different molecules, and these new substances have different properties from those of the reactants. • The total number of each type of atoms is conserved and thus the mass does not change.	**Energy and Matter** • Matter is conserved because atoms are conserved in physical and chemical processes.
Clarifications and Content Limits	**Clarification Statements** • Emphasize demonstrations of an understanding of the law of conservation of matter. • Emphasis is on law of conservation of matter and on physical models or drawings, including digital formats that represent atoms. • Models can include already balanced chemical equations. **Content Limits** • Assessment does not include the use of atomic masses, balancing symbolic equations, or intermolecular forces. • Assessment does not include stoichiometry or balancing equations. • Assessment is limited to simpler molecules, i.e., carbon dioxide, ammonia, sodium chloride, methanol, calcium chloride.		
Science Vocabulary Students Are Expected to Know	Transfer, molecule, element, conversion, phase change, dissolve, reactant, product.		
Science Vocabulary Students Are Not Expected to Know	Acid-base reactions, base, catalyst, reaction rate, endothermic/exothermic, equilibrium, oxidation-reduction reaction, chemical bond, electron sharing, electron transfer, ion, isotope.		
Phenomena			
Context/ Phenomena	Some example phenomena for MS-PS1-5: • An antacid tablet was added to water and bubbles appeared. The mass of the water and antacid tablet after the tablet dissolved was less than the mass of the water and tablet before they were mixed. • A strip of metal was added to acid in a test tube and a balloon was placed on top of the test tube. Bubbles appeared and after a few minutes, the balloon inflated.		

	• 100g of sugar completely dissolved in 100ml of water. After it dissolved, the mass of the mixture was 200g. • Steel wool was soaked in water ad left out to dry. The steel wool turned dark red, ad the mass of the steel wool after it dried was greater than before it was soaked in the water.
This Performance Expectation and associated Evidence Statements support the following Task Demands.	
Task Demands	
7. Select or identify from a collection of potential model components, including distractors, the components needed to model the phenomenon. Components might include atoms and molecules.	
8. Assemble or complete, from a collection of potential model components, an illustration or flow chart that is capable of representing the conservation of matter.*	
9. Manipulate the components of the model to demonstrate the changes, properties, processes, and/or events that act to result in the phenomenon.*	
10. Make predictions about the effects of changes in chemical reactions. Predictions can be made by manipulating model components, completing illustrations, or selecting from lists with distractors.**	
11. Identify missing components, relationships, or limitations of the model.	
12. Describe, select, or identify the relationships among components of a model that describe the conservation of matter, or explain the chemical reaction.	
13. Use the model to provide a causal account that matter is conserved during a chemical reaction by calculating the number of atoms or total mass of reactants and products.	

*denotes those task demands which are deemed appropriate for use in stand-alone item development
**TD4 may only be used in conjunction TD3

<table>
<tr><td>Performance Expectation</td><td colspan="3">MS-PS1-6
Undertake a design project to construct, test, and modify a device that either releases or absorbs thermal energy by chemical processes.</td></tr>
<tr><td>Dimensions</td><td>Constructing Explanations and Designing Solutions
• Undertake a design project, engaging in the design cycle, to construct and/or implement a solution that meets specific design criteria and constraints.</td><td>PS1.B: Chemical Reactions
• Some chemical reactions release energy, others store energy

ETS1.B: Developing Possible Solutions
• A solution needs to be tested, and then modified on the basis of the test results, in order to improve it. (secondary)

ETS1.C: Optimizing the Design Solution
• Although one design may not perform the best across all tests, identifying the characteristics of the design that performed the best in each test can provide useful information for the redesign process - that is, some of the characteristics may be incorporated into the new design. (secondary)
• The iterative process of testing the most promising solutions and modifying what is proposed on the basis of the test results leads to greater refinement and ultimately to an optimal solution. (secondary)</td><td>Energy and Matter
• The transfer of energy can be tracked as energy flows through a designed or natural system.</td></tr>
<tr><td>Clarifications and Content Limits</td><td colspan="3">Clarification Statements
• Emphasis is on the design, controlling the transfer of energy to the environment, and modification of a device using factors such as type and concentration of a substance.
• Examples of designs could involve chemical reactions such as dissolving ammonium chloride or calcium chloride.

Content Limits
• Assessment is limited to the criteria of amount, time, and temperature of substance in testing the device.
• <u>Students do not need to know:</u>
  ◦ Types of chemical reactions (decomposition, synthesis, single replacement, double replacement, combustion, etc.)
  ◦ How to balance a chemical equation</td></tr>
<tr><td>Science Vocabulary Students are Expected to Know</td><td colspan="3">Reactant, product, chemical bond, compound, molecule, solution, dissolve, solubility, concentration, chemical potential energy, thermal energy, system, environment, evaporate, condense</td></tr>
<tr><td>Science Vocabulary Students are Not Expected to Know</td><td colspan="3">Endothermic, exothermic, precipitant, solute, solvent, crystallization, dissolution, polar/polarity, ion, intermolecular force, intramolecular force, enthalpy, entropy, heat of solution, heat of reaction, microstates, equilibrium, saturate/saturation</td></tr>
<tr><td colspan="4" align="center">Phenomena</td></tr>
</table>

Context/ Phenomena	Engineering performance expecations are built around meaningful design problems rather than phenomena. For this performance expectation, the design problem and solutions replace phenomena. Some example design problems for MS-PS1-6: • Design a sport's injury pack that when used, will heat and soothe sore muscles. • Design a sport's injury cold pack that will help prevent swelling. • Design a self-heating pad that can warm ready-to-eat meals. • Design a device that can be used to keep electronics, like computers, from overheating.
This Performance Expectation and associated Evidence Statements support the following Task Demands.	
Task Demands	
1. Articulate, describe, illustrate, or select the relationships, interactions, and/or processes to be explained. This may entail sorting relevant from irrelevant information or features.	
2. Express or complete a causal chain explaining the chemical processes that resulted in the release or absorption of thermal energy. This may include indicating directions of causality in an incomplete model such as a flow chart or diagram, or completing cause and effect chains.	
3. Describe, identify, and/or select evidence supporting the inference of causation that is expressed in a causal chain and/or an explanation of the processes that cause the observed results.	
4. Use an explanation to predict the direction or the relative magnitude of a change in thermal energy of a chemical system, given a change in the amount/concentration of chemical substances in the system, the temperature of the substances in the system, and/or the amount of time the substances interact in the system.	
5. Identify or assemble from a collection, including distractors, the relevant aspects of the problem that given design solutions, if implemented, will resolve/improve.	
6. Using the given information, select or identify the criteria against which the device or solution should be judged.	
7. Using given data, propose, illustrate, or assemble a potential device (prototype) or solution.	
8. Using a simulator, test a proposed prototype and evaluate the outcomes, potentially including proposing and testing modifications to the prototype.	

<table>
<tr><td>Performance Expectation</td><td colspan="3">MS-PS2-1
Apply Newton’s Third Law to design a solution to a problem involving the motion of two colliding objects.</td></tr>
<tr><td>Dimensions</td><td>Constructing Explanations and Designing Solutions
• Apply scientific ideas or principles to design an object, tool, process, or system.</td><td>PS2.A: Forces and Motion
• For any pair of interacting objects, the force exerted by the first object on the second object is equal in strength to the force that the second object exerts on the first, but in the opposite direction (Newton’s Third Law).</td><td>Systems and System Models
• Models can be used to represent systems and their interactions—such as inputs, processes, and outputs—and energy and matter flows within systems.</td></tr>
<tr><td>Clarifications and Content Limits</td><td colspan="3">Clarification Statements
• Examples of practical problems could include the impact of collisions between two cars, between a car and stationary objects, and between a meteor and a space vehicle.

Content Limits
• Assessment is limited to vertical or horizontal interactions in one dimension.
• <u>Students do not need to know</u>: vector addition</td></tr>
<tr><td>Science Vocabulary Students are Expected to Know</td><td colspan="3">Conservation of momentum, energy transfer, transfer, force, balanced force, friction, direction of a force, impact, net force, inertia, action/reaction, gravity, acceleration, Newton, thrust, lift.</td></tr>
<tr><td>Science Vocabulary Students are Not Expected to Know</td><td colspan="3">Elastic collision, inelastic collision, impulse, coefficient of restitution, drag force, terminal velocity, friction coefficient, horizontal and vertical velocities (arc), aerodynamics, magnitude, vector.</td></tr>
<tr><td colspan="4">Phenomena</td></tr>
<tr><td>Context/ Phenomena</td><td colspan="3">Engineering performance expectations are built around meaningful design problems rather than phenomena. In this case, the design problems involve two colliding objects in a system. For this performance expectation, the design problem and competing solutions replace phenomena.

Some example design problems for MS-PS2-1:
• Testing different balls/objects for elementary students to throw at a dunk-tank target.
• Design a bike helmet that will keep the rider safe during a collision.
• Design a container that will protect vaccines from breaking as they are transported across rough terrain.
• Use Newton’s third law to create a system that will allow a ball to bounce higher than the height from which it was dropped.</td></tr>
<tr><td colspan="4">This Performance Expectation and associated Evidence Statements support the following Task Demands.</td></tr>
<tr><td colspan="4">Task Demands</td></tr>
<tr><td colspan="4">1. Identify or assemble from a collection, including distractors, the relevant aspects of the problem that given design solutions, if implemented, will resolve/improve.</td></tr>
<tr><td colspan="4">2. Using given information, select or identify constraints that the device or solution must meet, including cost, mass, and speed of objects and materials.</td></tr>
</table>

3. Using the given information, select or identify the criteria against which the device or solution should be judged.
4. Using given data, propose/illustrate/assemble a potential device (prototype) or solution.*
5. Using a simulator, test a proposed prototype and evaluate the outcomes; potentially propose and test modifications to the prototype.*

*denotes those task demands which are deemed appropriate for use in stand-alone item development

<table>
<tr><td>Performance Expectation</td><td colspan="3">MS-PS2-2
Plan an investigation to provide evidence that the change in an object’s motion depends on the sum of the forces on the object and the mass of the object.</td></tr>
<tr><td>Dimensions</td><td>Planning and Carrying Out Investigations
• Plan an investigation individually and collaboratively, and in the design: identify independent and dependent variables and controls, what tools are needed to do the gathering, how measurements will be recorded, and how many data are needed to support a claim</td><td>PS2.A: Forces and Motion
• The motion of an object is determined by the sum of the forces acting on it; if the total force on the object is not zero, its motion will change. The greater the mass of the object, the greater the force needed to achieve the same change in motion. For any given object, a larger force causes a larger change in motion.
• All positions of objects and the directions of forces and motions must be described in an arbitrarily chosen reference frame and arbitrarily chosen units of size. In order to share information with other people, these choices must also be shared.</td><td>Stability and Change
• Explanations of stability and change in natural or designed systems can be constructed by examining the changes over time and forces at different scales.</td></tr>
<tr><td>Clarifications and Content Limits</td><td colspan="3">Clarification Statements
• Emphasis is on:
  ◦ Balanced (Newton’s First law) and unbalanced forces in a system
  ◦ Qualitative comparisons of forces, masses and changes in motion (Newton’s Second Law)
  ◦ Frame of reference and specification of units

Content Limits
• Assessment is limited to forces and changes in motion in one-dimension in an inertial reference frame and to change in one variable at a time.
• <u>Students do not need to know:</u> trigonometry</td></tr>
<tr><td>Science Vocabulary Students are Expected to Know</td><td colspan="3">Applied force, balanced force, collision, force, unbalanced force, position over time, net force</td></tr>
<tr><td>Science Vocabulary Students are Not Expected to Know</td><td colspan="3">Newton’s Laws of Motion, acceleration, velocity, inertial frame of reference, momentum, friction</td></tr>
<tr><td colspan="4">Phenomena</td></tr>
<tr><td>Context/ Phenomena</td><td colspan="3">Some example phenomena for MS-PS2-2:
• A tennis ball is dropped on a trampoline and bounces up to a height, h. A bowling ball is then dropped on the same trampoline. The bowling ball bounces up to a height higher than h.
• A bowling ball is rolled towards a bowling pin. When the bowling ball hits the pin, the pin falls down. Then, a marble is rolled towards a bowling pin. When the marble hits the pin, the pin does not fall down.</td></tr>
</table>

	• A soccer player kicks the ball 50 yards. She then kicks another ball and it only goes 30 yards. • Two magnets of the same size are held apart from each other. One magnet is let go and moves towards the stationary magnet. When two other magnets are close to each other and one is let go, it moves toward the stationary magnet, faster.

This Performance Expectation and associated Evidence Statements support the following Task Demands.
Task Demands
1. Identify from a list, including distractors, the materials/tools needed for an investigation of how the change in an object's motion depends on the sum of the forces on the object and the mass of the object.
2. Identify the outcome data that should be collected in an investigation of how the sum of the forces on an object, as well as the object's mass, affect the change in motion of the object.
3. Evaluate the sufficiency and limitations of data collected to explain the phenomenon.
4. Make and/or record observations about how the sum of the forces on an object, and the mass of the object, affect the change in motion of the object.
5. Interpret and/or communicate the data from an investigation on how the change in motion of an object is affected by the sum of all forces and the mass of the object.
6. Explain or describe the causal processes that lead to the data that is observed in an investigation of how the forces on an object, and its mass, affect its change in motion.
7. Select, describe, or illustrate a prediction made by applying the findings from an investigation on how the forces on an object, and its mass, affect its change in motion.

<table>
<tr><td>Performance Expectation</td><td colspan="3">MS-PS2-3
Ask questions about data to determine the factors that affect the strength of electrical and magnetic forces.</td></tr>
<tr><td>Dimensions</td><td>Asking Questions and Defining Problems
• Ask questions that can be investigated within the scope of the classroom, outdoor environment, and museums and other public facilities with available resources and, when appropriate, frame a hypothesis based on observations and scientific principles.</td><td>PS2.B: Types of Interactions
• Electric and magnetic (electromagnetic) forces can be attractive or repulsive, and their sizes depend on the magnitudes of the charges, currents, or magnetic strengths involved and on the distances between the interacting objects.</td><td>Cause and Effect
• Cause and effect relationships may be used to predict phenomena in natural or designed systems.</td></tr>
<tr><td>Clarifications and Content Limits</td><td colspan="3">Clarification Statements
• Examples could include electromagnets, electric motors, or generators.
• Examples of data could include the effect of the number or turns of wire on the strength of an electromagnet, or of increasing the number or strength of magnets on the speed of an electric motor.

Content Limits
• Assessment about questions that require quantitative answers is limited to proportional reasoning and algebraic thinking.</td></tr>
<tr><td>Science Vocabulary Students are Expected to Know</td><td colspan="3">Attraction, conductor, electric current, electric field, electromagnetic field, electromagnet, frequency, induction, insulator, permanent magnet, polarity, repulsion, resistance, voltage.</td></tr>
<tr><td>Science Vocabulary Students are Not Expected to Know</td><td colspan="3">Lorentz force, electric potential, electromotive force.</td></tr>
<tr><td colspan="4">Phenomena</td></tr>
<tr><td>Context/ Phenomena</td><td colspan="3">Some example phenomena for MS-PS2-3:
• A radio emits music from its speakers. After a magnet in the speakers is removed, no sound can be heard.
• More electrical current is produced by a windmill when the wind speed is greater.
• Merchandise from a store that uses electromagnetic anti-shoplifting devices will set off an alarm at the exit if the tag is not removed.
• An electromagnet at a junkyard can lift old cars, while a homemade electromagnet cannot pick up much more than a few paper clips.</td></tr>
<tr><td colspan="4">This Performance Expectation and associated Evidence Statements support the following Task Demands.</td></tr>
<tr><td colspan="4">Task Demands</td></tr>
<tr><td colspan="4">1. Make and/or record observations about the factors that affect electromagnets, electric motors, or generators.</td></tr>
<tr><td colspan="4">2. Organize and/or arrange (e.g., using illustrations and/or labels), or summarize data to highlight trends, patterns, or correlations in the change in the strength of electrical and magnetic forces.</td></tr>
</table>

3. Generate or construct graphs, tables, or assemblages of illustrations and/or labels of data that document patterns, trends, or correlations in the factors that affect the strength of electric and magnetic forces. This may include sorting out distractors.
4. Explain or describe the causal processes that lead to the change in the strength of electrical and magnetic forces.
5. Use relationships identified in the data to predict the strength of electric and/or magnetic forces.
6. Select from a list of questions, including distractors, a scientifically testable question about factors that affect the strength of electrical or magnetic forces.

<table>
<tr><td>Performance Expectation</td><td colspan="3">MS-PS2-4
Construct and present arguments using evidence to support the claim that gravitational interactions are attractive and depend on the masses of interacting objects.</td></tr>
<tr><td>Dimensions</td><td>Engaging in Argument from Evidence
• Construct and present oral and written arguments supported by empirical evidence and scientific reasoning to support or refute an explanation or a model for a phenomenon or a solution to a problem.</td><td>PS2.B Types of Interactions
• Gravitational forces are always attractive. There is a gravitational force between any two masses, but it is very small except when one or both of the objects have large mass—e.g., Earth and the sun.</td><td>Systems and System Models
• Models can be used to represent systems and their interactions—such as inputs, processes, and outputs—and energy and matter flows within systems.</td></tr>
<tr><td>Clarifications and Content Limits</td><td colspan="3">Clarification Statements
• Examples of evidence of arguments could include data generated from simulations or digital tools, and charts displaying mass, strength of interaction, distance from the sun, and orbital periods of objects within the solar system.

Content Limits
• Assessment does not include Newton’s law of gravitation or Kepler’s laws.
• <u>Students do not need to know:</u> mathematical representations of gravity (values, units, etc.).</td></tr>
<tr><td>Science Vocabulary Students are Expected to Know</td><td colspan="3">Orbit, magnitude, galaxy, solar system, satellite, force fields, ellipse, proportional, period.</td></tr>
<tr><td>Science Vocabulary Students are Not Expected to Know</td><td colspan="3">Terminal velocity, relativity, gravitational energy, gravitational field, inverse square law.</td></tr>
<tr><td colspan="4">Phenomena</td></tr>
<tr><td>Context/ Phenomena</td><td colspan="3">Some example phenomena for MS-PS2-4:
• The moon orbits Earth.
• Astronauts fall more slowly when jumping on the moon than on Earth.
• A dropped apple falls toward Earth, but not toward the moon.
• Rockets have to travel extremely fast when they leave Earth.</td></tr>
<tr><td colspan="4">This Performance Expectation and associated Evidence Statements support the following Task Demands.</td></tr>
<tr><td colspan="4">Task Demands</td></tr>
<tr><td colspan="4">1. Articulate, describe, illustrate, or select the relationships, interactions, and/or processes to be explained. This may entail sorting relevant from irrelevant information.</td></tr>
<tr><td colspan="4">2. Predict outcomes when properties or proximity of the objects are changed, given the inferred cause and effect relationships</td></tr>
<tr><td colspan="4">3. Describe, identify, and/or select information needed to support an explanation.**</td></tr>
</table>

4. Identify patterns or evidence in the data that support conclusions about the relationship between mass and gravity.*
5. Using evidence, explain the relationship between mass and gravity.*

*denotes those task demands which are deemed appropriate for use in stand-alone item development
**TD3 may be used only in conjunction with TD4 or TD5.

<table>
<tr><td>Performance Expectation</td><td colspan="3">MS-PS2-5
Conduct an investigation and evaluate the experimental design to provide evidence that fields exist between objects exerting forces on each other even though the objects are not in contact.</td></tr>
<tr><td>Dimensions</td><td>Planning and Carrying Out Investigations
• Conduct an investigation and evaluate the experimental design to produce data to serve as the basis for evidence that can meet the goals of the investigation.</td><td>PS2.B: Types of Interactions
• Forces that act at a distance (electric, magnetic, and gravitational) can be explained by fields that extend through space and can be mapped by their effect on a test object (a charged object, or a ball, respectively).</td><td>Cause and Effect
• Cause and effect relationships may be used to predict phenomena in natural or designed systems.</td></tr>
<tr><td>Clarifications and Content Limits</td><td colspan="3">Clarification Statements
• Examples of this phenomenon could include the interactions of magnets, electrically-charged strips of tape, and electrically-charged pith balls.
• Examples of investigations could include first-hand experiences or simulations.

Content Limits
• Assessment is limited to electric and magnetic fields, and limited to qualitative evidence for the existence of fields.</td></tr>
<tr><td>Science Vocabulary Students are Expected to Know</td><td colspan="3">Conductor, electric charge, electric current, electric force, electromagnetic field, electromagnet, frequency, induction, insulator, magnetic field lines, magnetic force, permanent magnet, polarity, repulsion, resistance, voltage, direction, magnitude, ampere, charged particle, volts, gravity</td></tr>
<tr><td>Science Vocabulary Students are Not Expected to Know</td><td colspan="3">Lorentz force, electric potential, electromotive force, permeating, vector field, quantum property, Laplace force, Right-hand rule, Ampere’s Law, electrodynamics, magnetic dipole, Coulomb force, electrostatic, general relativity</td></tr>
<tr><td colspan="4">Phenomena</td></tr>
<tr><td>Context/ Phenomena</td><td colspan="3">Some example phenomena for MS-PS2-5:
• A compass is opened and set on a table. The needle spins for a bit and then settles pointing north.
• Two blue-painted metal boxes sit on a table. With a pocket knife, a person easily scratches some of the paint off of one box. But they cannot remove the paint from the other box.
• A person walks across a carpeted floor in stocking feet. They touch another person who is sitting in a chair, delivering a large shock.
• A multimeter records the presence of an electric current when a coil rotates near a magnet.</td></tr>
<tr><td colspan="4">This Performance Expectation and associated Evidence Statements support the following Task Demands.</td></tr>
<tr><td colspan="4">Task Demands</td></tr>
<tr><td colspan="4">1. Identify from a list, including distractors, the materials/tools/steps needed for an investigation of fields that exist between objects exerting forces on each other even though the objects are not in contact.</td></tr>
<tr><td colspan="4">2. Identify the outcome data that should be collected for a given purpose in an investigation of fields that exist between objects exerting forces on each other even though the objects are not in contact.</td></tr>
</table>

3. Evaluate the sufficiency and limitations of data collected to explain the phenomenon.
4. Make and/or record observations about fields that exist between objects exerting forces on each other even though the objects are not in contact.
5. Interpret and/or communicate the data from an investigation of the field that exists between two objects exerting forces on each other even though the objects are not in contact.
6. Explain, describe, or identify the causal processes that lead to the observed data about the field that exists between two objects exerting forces on each other even though the objects are not in contact.
7. Select, describe, or illustrate a prediction made by applying the findings from an investigation of the field that exists between two objects exerting forces on each other even though the objects are not in contact

<table>
<tr><td>Performance Expectation</td><td colspan="3">MS-PS3-1
Construct and interpret graphical displays of data to describe the relationships of kinetic energy to the mass of an object and to the speed of an object</td></tr>
<tr><td>Dimensions</td><td>Analyzing and Interpreting Data
• Construct and interpret graphical displays of data to identify linear and nonlinear relationships</td><td>PS3.A: Definitions of Energy
• Motion energy is properly called kinetic energy it is proportional to the mass of the moving object and grows with the square of its speed.</td><td>Scale, Proportion, and Quantity
• Proportional relationships (e.g. speed as the ratio of distance traveled to time taken) among different types of quantities provide information about the magnitude of properties and processes.</td></tr>
<tr><td>Clarifications and Content Limits</td><td colspan="3">Clarification Statements
• Emphasis is on descriptive relationships between kinetic energy and mass separately from kinetic energy and speed.
• Examples could include:
    • Riding a bicycle at different speeds
    • Rolling different sizes of rocks downhill
    • Getting hit by a wiffle ball vs a tennis ball

Content Limits
• <u>Students do not need to know:</u> vectors such as velocity, the exact formula for the kinetic energy of an object or how to make calculations using the formula.</td></tr>
<tr><td>Science Vocabulary Students are Expected to Know</td><td colspan="3">Forms of energy, magnitude, motion energy, proportional, ratio, square root, potential energy, elastic collision, inelastic collision</td></tr>
<tr><td>Science Vocabulary Students are Not Expected to Know</td><td colspan="3">Velocity, vector, inertial frame of reference, acceleration, deceleration, relative motion, Newtonian Mechanics</td></tr>
<tr><td colspan="4">Phenomena</td></tr>
<tr><td>Context/ Phenomena</td><td colspan="3">Some example phenomena for MS-PS3-1:
• Balls of different masses are dropped into a pile of snow. A graph of the mass vs. the depth of the indent is shown.
• A pendulum is dropped so that it hits a box on the ground. A graph of the drop height vs the distance the box travels is shown.
• A ball thrown at a wall will bounce back a certain distance. A table of the speed of the ball vs. the distance it bounces back is given.
• Trains with differing amounts of train cars all come to a stop. A table of the number of train cars vs stopping distance is given.</td></tr>
<tr><td colspan="4">This Performance Expectation and associated Evidence Statements support the following Task Demands.</td></tr>
<tr><td colspan="4">Task Demands</td></tr>
<tr><td colspan="4">1. Organize and/or arrange (e.g., using illustrations and/or labels), or summarize data to highlight trends, patterns, or correlations among observations and data concerning the mass, speed and kinetic energy of objects. This may include sorting out distractors.</td></tr>
</table>

2. Generate/construct graphs, tables, or assemblages of illustrations and/or labels of data that document patterns, trends, or correlations in how the kinetic energy of an object changes with its mass and its speed. This may include sorting out distractors.
3. Use relationships identified in the data to predict how the kinetic energy of an object will change based on a change in speed of the object or mass of the object.
4. Identify patterns or evidence in the data that supports inferences about how kinetic energy changes with the speed of an object and the mass of an object.

<table>
<tr><td>Performance Expectation</td><td colspan="3">MS-PS3-2
Develop a model to describe that when the arrangement of objects interacting at a distance changes, different amounts of potential energy are stored in the system.</td></tr>
<tr><td>Dimensions</td><td>Developing and Using Models
• Develop a model to describe unobservable mechanisms.</td><td>PS3.A: Definitions of Energy
• A system of objects may also contain stored (potential) energy, depending on their relative positions.

PS3.C: Relationship Between Energy and Forces
• When two objects interact, each one exerts a force on the other that can cause energy to be transferred to or from the object.</td><td>Systems and System Models
• Models can be used to represent systems and their interactions—such as inputs, processes, and outputs—and energy and matter flows within systems.</td></tr>
<tr><td>Clarifications and Content Limits</td><td colspan="3">Clarification Statements
• Emphasis is on relative amounts of potential energy, not on calculations of potential energy.
• Examples of objects within systems interacting at varying distances could include the Earth and either a roller coaster cart at varying positions on a hill or objects at varying heights on shelves, changing the direction/orientation of a magnet, and a balloon with static electrical charge being brought closer to a classmate's hair.
• Examples of models could include representations, diagrams, pictures, and written descriptions of systems.

Content Limits
• Assessment does not include calculations of kinetic and potential energy.
• Assessment is limited to two objects and electric, magnetic, and gravitational interactions.</td></tr>
<tr><td>Science Vocabulary Students are Expected to Know</td><td colspan="3">Electron, proton, distribution of charged particles, electrical charge, negatively charged, positively charged, neutrally charged, magnetic polarity, conductor, insulator, electromagnet.</td></tr>
<tr><td>Science Vocabulary Students are Not Expected to Know</td><td colspan="3">Oscillation, harmonic oscillator, period, momentum, spring constant, equilibrium position, acceleration of gravity, work, power, mechanical advantage, Work-energy theorem, rotational motion, translational motion, torque, moment, Coulomb's law, Faraday cage, triboelectricity, electric potential, gravitational potential.</td></tr>
<tr><td colspan="4" align="center">Phenomena</td></tr>
<tr><td>Context/ Phenomena</td><td colspan="3">Some example phenomena for MS-PS3-2:
• A roller coaster track contains two hills of equal size. A roller coaster car sitting on the first hill is released and allowed to roll down the tracks of the first hill. The car comes to a stop before it reaches the top of the second hill.
• Two wrecking ball cranes sit next to two concrete buildings. Crane A has a ball that has less mass than the ball of Crane B. Both cranes swing their balls toward the buildings. Crane A's ball starts out higher than Crane B's ball. Crane A's ball does substantially more damage to the building than Crane B's ball.
• The poles of an electromagnet can be reversed by reversing the electromagnet's connection to a battery.
• An empty shopping cart rolls down a hill in a parking lot and dents a parked car, while a full shopping cart rolls across a flat lot and does not damage a parked car.</td></tr>
</table>

This Performance Expectation and associated Evidence Statements support the following Task Demands.
Task Demands
1. Select or identify from a collection of potential model components, including distractors, the components needed to model different amounts of potential energy stored in a system, compared to the distance between interacting objects. Components might include: energy source, objects in motion, and boundaries of system.
2. Assemble or complete, from a collection of potential model components, an illustration or flow chart that is capable of representing changes in potential energy stored in a system. This *does not* include labeling an existing diagram.
3. Manipulate the components of the model to demonstrate the changes, properties, processes, and/or events that act to result in the changes in potential energy.
4. Make predictions about the effects of changes in distances between interacting objects and the potential energy stored in the system. Predictions can be made by manipulating model components, completing illustrations, or selecting from lists with distractors.
5. Given models or diagrams of a system containing potential energy, identify how the energy changes over time in a given scenario OR identify the properties of the variables that cause the changes.
6. Identify missing components, relationships, or other limitations of the model.
7. Describe, select, or identify the relationships among components of a model that describe changes in potential energy of a system when the distance between interacting objects changes.

<table>
<tr><td>Performance Expectation</td><td colspan="3">MS PS3-3
Apply scientific principles to design, construct, and test a device that either minimizes or maximizes thermal energy transfer.</td></tr>
<tr><td>Dimensions</td><td>Constructing Explanations and Designing Solutions
• Apply scientific ideas or principles to design, construct, and test a design of an object, tool, process or system.</td><td>PS3.A: Definitions of Energy
• Temperature is a measure of the average kinetic energy of particles of matter. The relationship between the temperature and the total energy of a system depends on the types, states, and amounts of matter present.

PS3.B: Conservation of Energy
• Energy is spontaneously transferred out of hotter regions or objects and into colder ones

ETS1.A: Defining and Delimiting an Engineering Problem
• The more precisely a design task’s criteria and constraints can be defined, the more likely it is that the designed solution will be successful. Specification of constraints includes consideration of scientific principles and other relevant knowledge that is likely to limit possible solutions.

ETS1.B: Developing Possible Solutions
• A solution needs to be tested, and then modified on the basis of the test results, in order to improve it. There are systematic processes for evaluating solutions with respect to how well they meet criteria and constraints of a problem.</td><td>Energy and Matter
• The transfer of energy can be tracked as energy flows through a designed or natural system.</td></tr>
<tr><td>Clarifications and Content Limits</td><td colspan="3">Clarification Statement
• Examples of devices could include an insulated box and a Styrofoam cup.

Content Limits
• Students should be given the problem to solve.
• <u>Students do not need to know</u>: Calculate energy of the system or change in energy.</td></tr>
<tr><td>Science Vocabulary Students are Expected to Know</td><td colspan="3">Temperature, kinetic energy, energy transfer, conductor, insulator, convection conduction and radiation.</td></tr>
<tr><td>Science Vocabulary Students are Not Expected to Know</td><td colspan="3">Energy units (joules, amperes), charged particles</td></tr>
<tr><td colspan="4">Phenomena</td></tr>
<tr><td>Context/ Phenomena</td><td colspan="3">Engineering performance expectations are built around meaningful design problems rather than phenomena. For this performance expectation, design problems or design solutions replace phenomena.

Some examples of design problems for MS-PS3-3:
• A heated swimming pool needs to be covered to reduce energy costs in the winter.</td></tr>
</table>

	• Many cooks prefer pans that heat more evenly. Which materials should pans be made of? • Design a more energy-efficient window. • Choose the materials for a pot holder.
This Performance Expectation and associated Evidence Statements support the following Task Demands.	
Task Demands	
1. Identify or assemble from a collection the relevant aspects of the problem that given design solutions for either minimizing or maximizing thermal energy transfer, if implemented, will resolve/improve.	
2. Using the given information, select or identify the criteria against which the device or solution that either minimizes or maximizes thermal energy transfer should be judged.	
3. Using given information, select or identify constraints that the device or solution that either minimizes or maximizes thermal energy transfer must meet.	
4. Using given data, propose, illustrate, and/or assemble a potential device (prototype) or solution that either minimizes or maximizes thermal energy transfer.	
5. Using a simulator, test a proposed prototype and evaluate the outcomes, potentially including proposing and testing modifications to the prototype.	

<table>
<tr><td>Performance Expectation</td><td colspan="3">MS-PS3-4
Plan an investigation to determine the relationships among energy transferred, type of matter, mass, and change in the average kinetic energy of particles, as measured by the temperature of a sample.</td></tr>
<tr><td>Dimensions</td><td>Planning and Carrying Out Investigations
• Plan an investigation individually and collaboratively and, in the design, identify independent and dependent variables and controls, what tools are needed to do the gathering, how measurements will be recorded, and how many data are needed to support a claim.</td><td>PS3.A: Definitions of Energy
• Temperature is a measure of the average kinetic energy of particles of matter. The relationship between the temperature and the total energy of a system depends on the types, states, and amounts of matter present.

PS3.B: Conservation of Energy and Energy Transfer
• The amount of energy transfer needed to change the temperature of a sample of matter by a given amount depends on the nature of the matter, the size of the sample, and the environment.</td><td>Scale, Proportion, and Quantity
• Proportional relationships (e.g., speed as the ratio of distance traveled to time taken) among different types of quantities provide information about the magnitude of properties and processes.</td></tr>
<tr><td>Clarifications and Content Limits</td><td colspan="3">Clarification Statements
• Examples of experiments could include comparing final water temperatures after different masses of ice melted in the same volume of water with the same initial temperature; the temperature change of samples of different materials with the same mass as they cool or heat in the environment; or the same material with different masses when a specific amount of energy is added.

Content Limits
• Assessment does not include calculating the total amount of thermal energy transferred.</td></tr>
<tr><td>Science Vocabulary Students are Expected to Know</td><td colspan="3">Volume, collide, collision, heat conduction, particle, stored energy, transfer, average, proportional, ratio, thermal energy</td></tr>
<tr><td>Science Vocabulary Students are Not Expected to Know</td><td colspan="3">Thermal equilibrium, thermodynamics</td></tr>
<tr><td colspan="4">Phenomena</td></tr>
<tr><td>Context/ Phenomena</td><td colspan="3">Some example phenomena for MS-PS3-4:
• A mug of hot coffee is set on a cork coaster. After letting the mug of coffee sit for a while, a person picks up the mug and the coaster and notices that both the mug and coaster are warm.
• When placed over the same heat source, water takes longer to reach $100C^{o}$ than a cola soft drink.
• Pot holders work well when they’re dry. When they’re wet, they don’t.
• A metal spoon used to stir a hot beverage gets hot much more quickly than a wooden spoon.</td></tr>
<tr><td colspan="4">This Performance Expectation and associated Evidence Statements support the following Task Demands.</td></tr>
</table>

Task Demands
1. Identify from a list, including distractors, the materials/tools needed for an investigation of how thermal energy is transferred to and from the environment and to and from materials of different/ same types of matter and different/ same masses.
2. Identify the data that should be collected in an investigation of how thermal energy is transferred to and from the environment and to and from materials of different/ same types of matter and different/ same masses.
3. Evaluate the sufficiency and limitations of data collected to explain a phenomenon.
4. Make and/or record observations about time, mass of materials, type of materials, initial and final average kinetic energy (temperature) of materials, and the surrounding environment.
5. Interpret and/or communicate data from an investigation.
6. Explain or describe the causal processes that lead to observed data.
7. Select, describe, or illustrate a prediction made by applying the findings from an investigation.
8. Assemble or specify a controlled experiment or investigation to evaluate the effect of the type of matter, amount of heat, or volume of material heated.

<table>
<tr><td>Performance Expectation</td><td colspan="3">MS-PS3-5
Construct, use, and present arguments to support the claim that when the kinetic energy of an object changes, energy is transferred to or from the object.</td></tr>
<tr><td>Dimensions</td><td>Engaging in Argument from Evidence
• Construct, use, and present oral and written arguments supported by empirical evidence and scientific reasoning to support or refute an explanation or a model for a phenomenon.</td><td>PS3.B: Conservation of Energy and Energy Transfer
• When the kinetic energy of an object changes, there is inevitably some other change in energy at the same time.</td><td>Energy and Matter
• Energy may take different forms (e.g., energy in fields, thermal energy, and energy of motion).</td></tr>
<tr><td>Clarifications and Content Limits</td><td colspan="3">Clarification Statements
• Emphasis is on understanding that when the kinetic energy of an object increases or decreases, the energy (e.g., kinetic, thermal, potential, light, sound) of other objects or the surroundings within the system increases or decreases, indicating that energy was transferred to or from the object.
• Emphasis is on knowing that temperature is the measure of the average kinetic energy of particles of matter.
• Examples of empirical evidence used in arguments could include an inventory or other representation of the energy before and after the transfer in the form of temperature changes or motion of an object.

Content Limits
• Assessment does not include any calculations of energy or energy flow.</td></tr>
<tr><td>Science Vocabulary Students Are Expected to Know</td><td colspan="3">Potential energy, heat energy, closed system, open system, friction, joule, force, transformation of energy, thermometer, Fahrenheit, Celsius, pendulum, sound energy, conservation of energy</td></tr>
<tr><td>Science Vocabulary Students Are Not Expected to Know</td><td colspan="3">Co-efficient of kinetic energy, air resistance, work, energy efficiency, chemical energy, electrical energy, machine (for transforming energy), mechanical energy.</td></tr>
<tr><td colspan="4">Phenomena</td></tr>
<tr><td>Context/ Phenomena</td><td colspan="3">Some example phenomena for MS-PS3-5:
• The Riverside geyser in the Upper Geyser Basin at Yellowstone National Park throws out jets of hot water into the air at regular intervals.
• When the brakes are applied, sparks fly out between the wheels and the metal tracks as a train slows down.
• Bowling pins fall over and start to roll when struck by a bowling ball.
• A hot air balloon lifts off the ground as the burner is lit under the balloon.</td></tr>
<tr><td colspan="4">This Performance Expectation and associated Evidence Statements support the following Task Demands.</td></tr>
<tr><td colspan="4">Task Demands</td></tr>
<tr><td colspan="4">1. Articulate, describe, illustrate, or select the relationships, interactions, and/or processes to be explained. This may entail sorting relevant from irrelevant information.</td></tr>
</table>

2. Predict outcomes when the kinetic energy of an object changes, given the inferred cause and effect relationships.
3. Describe, identify, and/or select information needed to support an explanation of a change in kinetic energy or energy transfer.
4. Identify patterns or evidence in the data that support the claim that the kinetic energy of an object changes as energy is transferred to or from the object.
5. Using evidence, explain the relationship between the kinetic energy of an object and changes to the object or the surroundings, as energy is transferred to or from the object.
6. Manipulate the components of a model to demonstrate that the kinetic energy of an object changes as energy is transferred to or from the object.

<table>
<tr><td>Performance Expectation</td><td colspan="3">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.</td></tr>
<tr><td>Dimensions</td><td>Using Mathematics and Computational Thinking
• Use mathematical representations to describe and/or support scientific conclusions and design solutions.</td><td>PS4.A: Wave Properties
• A simple wave has a repeating pattern with a specific wavelength, frequency, and amplitude.</td><td>Patterns
• Graphs and charts can be used to identify patterns in data.</td></tr>
<tr><td>Clarifications and Content Limits</td><td colspan="3">Clarification Statements
• Emphasize describing waves with both quantitative and qualitative thinking.
• Examples could include using graphs, charts, computer simulations, or physical models to demonstrate amplitude and energy correlation.
• All equations and formulas must be provided and be age-appropriate.

Content Limits
• Assessment does not include electromagnetic waves and is limited to standard repeating waves.
• Assessment does not include identifying or knowing characteristics of different types of waves (mechanical, electromagnetic, sonic, etc.).
• <u>Students do not need to know</u>: how two waves carrying the same energy can have different amplitudes when introduced into materials of different densities and elasticities.</td></tr>
<tr><td>Science Vocabulary Students Are Expected to Know</td><td colspan="3">Speed, force, kinetic energy, proportional, sound wave, wavelength, frequency, resting position, medium, crest, trough</td></tr>
<tr><td>Science Vocabulary Students Are Not Expected to Know</td><td colspan="3">Elastic, seismic wave, oscillate.</td></tr>
<tr><td colspan="4">Phenomena</td></tr>
<tr><td>Context/ Phenomena</td><td colspan="3">Some example phenomena for MS-PS4-1:
• The 1896 Sanriku earthquake off the coast of Japan generated ocean waves that reached a height of 100 feet (30 m).
• Compared to a megaphone that sends sound messages up to 300 meters away, a Long Range Acoustic Device (LRAD) sends messages that can be heard up to 5,500 meters away.
• Scientists at the Swiss Federal Institute in Zurich caused a toothpick to levitate using sound waves.
• A wave travels down a rope from one student to another when the first student shakes it.</td></tr>
<tr><td colspan="4">This Performance Expectation and associated Evidence Statements support the following Task Demands.</td></tr>
<tr><td colspan="4">Task Demands</td></tr>
<tr><td colspan="4">1. Compile and analyze data to make an inference about the relationship between amplitude and energy of a wave. This may include sorting out relevant from irrelevant data in the given information.</td></tr>
<tr><td colspan="4">2. Organize and/or arrange (e.g., using illustrations and/or labels) or summarize data to highlight trends, patterns, or correlations that reflect how energy changes with amplitude of a wave and vice versa.</td></tr>
</table>

3. Identify how wave characteristics correspond to physical observations (e.g., wave amplitude corresponds to sound volume).
4. Use relationships identified in the data to predict the energy or amplitude change of a wave if the other parameter is changed.
5. Based on data, calculate or estimate one property of a wave (energy or amplitude) and the relationships between different properties of a wave.
6. Use graphs, charts, simulations, or physical models to demonstrate amplitude and energy correlation.

<table>
<tr><td>Performance Expectation</td><td colspan="3">MS-PS4-2
Develop and/or use a model to describe that waves are reflected, absorbed, or transmitted through various materials.</td></tr>
<tr><td>Dimensions</td><td>Developing and Using Models
• Develop and/or use a model to predict and/or describe phenomena.</td><td>PS4.A: Wave Properties
• A sound wave needs a medium through which it is transmitted.

PS4.B: Electromagnetic Radiation
• When light shines on an object, it is reflected, absorbed, or transmitted through the object, depending on the object’s material and the frequency (color) of the light.
• The path that light travels can be traced as straight lines, except at surfaces between different transparent materials (e.g., air and water, air and glass), where the light path bends.
• A wave model of light is useful for explaining brightness, color, and the frequency-dependent bending of light at a surface between media.
• However, because light can travel through space, it cannot be a matter wave, like sound or water waves.</td><td>Structure and Function
• Structures can be designed to serve particular functions by taking into account properties of different materials, and how materials can be shaped and used.</td></tr>
<tr><td>Clarifications and Content Limits</td><td colspan="3">Clarification Statement
• Emphasis is on both light and mechanical waves.
• Examples of models could include drawings, simulations, and written descriptions. This includes amplitudes, frequencies, and wave lengths.

Content Limits
• Assessment is limited to qualitative applications pertaining to light and mechanical waves, not quantitative.
• Assessment does not include:
    • Particle movement and compression waves
    • Constructive or destructive interference</td></tr>
<tr><td>Science Vocabulary Students are Expected to Know</td><td colspan="3">Refracted, medium, transparent, frequency, brightness, color, bending, amplitude, sound wave, light wave, path, propagation, filter, barrier, lens, mirror, mechanical waves, electromagnetic, visible light, ray, prism, wavelength.</td></tr>
<tr><td>Science Vocabulary Students are Not Expected to Know</td><td colspan="3">Longitudinal wave, transverse wave, compression wave, seismic waves, radio wave, microwave, infrared, ultraviolet, x-rays, gamma rays, angle of incidence, concave, convex, diffraction, constructive interference, destructive interference</td></tr>
<tr><td colspan="4">Phenomena</td></tr>
<tr><td>Context/ Phenomena</td><td colspan="3">Some example phenomena for MS-PS4-2:
• One part of a straw appears to be broken from the rest of the straw when viewed through the side of a glass of water.
• Music placed near a lake can be heard clearly while sitting on the shore. However, while swimming under the water, the sound cannot be heard as clearly.</td></tr>
</table>

	• Objects are more visible during a moonlit night when there is snow on the ground vs. when there is no snow on the ground. • Loud music moves the leaves of a plant.
This Performance Expectation and associated Evidence Statements support the following Task Demands.	
Task Demands	
1. Select from a collection of potential model components including distractors, the components needed to model the phenomenon. Components might include type of wave, properties of the wave, the materials with which the waves interact, the position of the source of the wave, etc.	
2. Assemble, from a collection of potential model components, an illustration or flow chart that is capable of representing the movement, transmission, reflection, refraction, and absorption of waves. This does not include labeling an existing diagram.	
3. Manipulate the components of a model to demonstrate the changes that cause the observed phenomenon.	
4. Manipulate the components of a model to predict the behavior of waves in an alternate scenario.	
5. Given models or diagrams of how a wave interacts with different materials, identify the wave properties and how they change in each scenario OR identify the properties of the different materials that cause the wave to behave differently.	
6. Identify missing components, relationships, or other limitations of the model.	

<table>
<tr><td>Performance Expectation</td><td colspan="3">MS-PS4-3
Integrate qualitative scientific and technical information to support the claim that digitized signals are a more reliable way to encode and transmit information than analog signals.</td></tr>
<tr><td>Dimensions</td><td>Obtaining, Evaluating, and Communicating Information
• Integrate qualitative scientific and technical information in written text with information contained in media and visual displays to clarify claims and findings.</td><td>PS4.C: Information Technologies and Instrumentation
• Digitized signals (sent as wave pulses) are a more reliable way to encode and transmit information</td><td>Structure and Function
• Structures can be designed to serve particular functions.</td></tr>
<tr><td>Clarifications and Content Limits</td><td colspan="3">Clarification Statements
• Emphasis is on a basic understanding that waves can be used for communication purposes.
• Examples could include using fiber optic cable to transmit light pulses, radio wave pulses in wifi devices, and conversion of stored binary patterns to make sound or text on a computer screen.
• Examples could also include using vinyl record vs. digital song files, film cameras vs. digital cameras, or alcohol thermometers vs. digital thermometers.

Content Limits
• Assessment does not include binary counting.
• Assessment does not include the specific mechanism of any given device.
• <u>Students do not need to know</u>:
o Specifics about binary or any other coding process.
o How certain mechanisms work other than the fact that they are either analog or digital.
o Students are not responsible for knowing the different parts of mechanisms: hard drives, USB cables, flash drives, and servers.</td></tr>
<tr><td>Science Vocabulary Students Are Expected to Know</td><td colspan="3">Computer, machine, communicate, electricity, device, coded, decode, conversion/convert, digitize, encode, radio wave</td></tr>
<tr><td>Science Vocabulary Students Are Not Expected to Know</td><td colspan="3">Binary, emit, photoelectric, pixel, electromagnetic radiation, radiation, wave packet, wave source, ohm, photon, microwave, ultraviolet, volt, ampere.</td></tr>
<tr><td colspan="4">Phenomena</td></tr>
<tr><td>Context/ Phenomena</td><td colspan="3">Some example phenomena for Standard MS-PS4-3:
• A digital scale gives better precision on weight measurements than analog.
• Digital films are higher quality than analog films (from a film reel).
• Digital measurements provide precise values compared to analog measurements
• Digital data can be stored in a server and easily retrieved if the hardware breaks, while analog data are lost if the hardware is broken.</td></tr>
<tr><td colspan="4">This Performance Expectation and associated Evidence Statements support the following Task Demands.</td></tr>
<tr><td colspan="4">Task Demands</td></tr>
</table>

1. Identify evidence that is sufficient to support the claim that digital signals are a more reliable way to store and transmit information than analog signals.
2. Citing evidence, identify specific features of digital signals that make them more reliable than analog signals OR identify specific examples of how digitization of a certain technology has advanced science.
3. Gather, read and synthesize information from multiple sources and assess the credibility, accuracy, and possible bias of each publication; describe how they are supported or not supported by evidence.
4. Evaluate data and/or conclusions in scientific and technical texts in light of competing information.

<table>
<tr><td>Performance Expectation</td><td colspan="3">MS-LS1-1
Conduct an investigation to provide evidence that living things are made of cells; either one cell or many different numbers and types of cells.</td></tr>
<tr><td>Dimensions</td><td>Planning and Carrying Out Investigations
• Conduct an investigation to produce data to serve as the basis for evidence that meets the goals of an investigation.</td><td>LS1.A: Structure and Function
• All living things are made up of cells, which is the smallest unit that can be said to be alive. An organism may consist of one single cell (unicellular) or many different numbers and types of cells (multicellular).</td><td>Scale, Proportion, and Quantity
• Phenomena that can be observed at one scale may not be observable at another scale.</td></tr>
<tr><td>Clarifications and Content Limits</td><td colspan="3">Clarification Statements
• Emphasis is on developing evidence that living things are made of cells, distinguishing between living and non-living things, and understanding that living things may be made of one cell or many varying cells.

Content Limits
• <u>Students do not need to know</u>:
○ The structures or functions of specific organelles or different proteins
○ Systems of specialized cells
○ The mechanisms by which cells are alive
○ Specifics of DNA and proteins or of cell growth and division
○ Endosymbiotic theory
○ Histological procedures.</td></tr>
<tr><td>Science Vocabulary Students are Expected to Know</td><td colspan="3">Multicellular, unicellular, tissues, organ, system, organism hierarchy, bacteria, colonies, yeast, prokaryote, eukaryote, magnify, microscope, DNA, nucleus, cell wall, cell membrane, algae, chloroplast(s), chromosomes</td></tr>
<tr><td>Science Vocabulary Students are Not Expected to Know</td><td colspan="3">Differentiation, mitosis, meiosis, genetics, cellular respiration, energy transfer, RNA, protozoa, amoeba, histology, Protista, archaea, nucleoid, plasmid, diatoms, cyanobacteria.</td></tr>
<tr><td colspan="4">Phenomena</td></tr>
<tr><td>Context/ Phenomena</td><td colspan="3">Some example phenomena for MS-LS1-1:
• Plant leaves and roots have tiny box-like structures that can be seen under a microscope.
• Small creatures can be seen swimming in samples of pond water viewed through a microscope.
• Different parts of a frog’s body (muscles, skin, tongue, etc.) are observed under a microscope, and are seen to be composed of cells.
• One-celled organisms (bacteria, protists) perform the eight necessary functions of life, but nothing smaller has been seen to do this.</td></tr>
<tr><td colspan="4">This Performance Expectation and associated Evidence Statements support the following Task Demands.</td></tr>
<tr><td colspan="4">Task Demands</td></tr>
<tr><td colspan="4">1. Identify from a list, including distractors, the materials/tools needed for an investigation to find the smallest unit of life (cell).</td></tr>
</table>

2. Identify the outcome data that should be collected in an investigation of the smallest unit of living things.
3. Evaluate the sufficiency and limitations of data collected to explain that the smallest unit of living things is the cell.
4. Make and/or record observations about whether the sample contains cells or not.*
5. Interpret and/or communicate data from the investigation to determine if a specimen is alive or not.
6. Construct a statement to describe the overall trend suggested by the observed data.

*denotes those task demands which are deemed appropriate for use in stand-alone item development

<table>
<tr><td>Performance Expectation</td><td colspan="3">MS-LS1-2
Develop and use a model to describe the function of a cell as a whole and ways the parts of cells contribute to the function.</td></tr>
<tr><td>Dimensions</td><td>Developing and Using Models
• Develop and use a model to describe phenomena.</td><td>LS1.A: Structure and Function
• Within cells, special structures are responsible for particular functions, and the cell membrane forms the boundary that controls what enters and leaves the cell.</td><td>Structure and Function
• Complex and microscopic structures and systems can be visualized, modeled, and used to describe how their function depends on the relationships among its parts; therefore, complex natural structures/systems can be analyzed to determine how they function.</td></tr>
<tr><td>Clarifications and Content Limits</td><td colspan="3">Clarification Statements
• Emphasize the cell functioning as a whole system and the primary role of identified parts of the cell, specifically the nucleus, chloroplasts, mitochondria, cell membrane, and cell wall.

Content Limits
• Assessment of organelle structure/function relationships is limited to the cell wall and cell membrane.
• Assessment of the function of the other organelles is limited to their relationship to the whole cell.
• Assessment does not include the biochemical function of cells or cell parts.
• <u>Students do not need to know</u>: protein synthesis, cell division (mitosis), reproduction (meiosis).</td></tr>
<tr><td>Science Vocabulary Students are Expected to Know</td><td colspan="3">Eukaryote, prokaryote, nucleus, chloroplast, mitochondrion, cell membrane, cell wall, diffusion, osmosis, photosynthesis, cellular respiration, sugar, DNA, RNA, energy, bacteria, cytoplasm, organelle.</td></tr>
<tr><td>Science Vocabulary Students are Not Expected to Know</td><td colspan="3">Golgi, ribosome, endoplasmic reticulum, enzyme, replication, mitosis, meiosis, glucose, chromosome, protein channels, lysosome, vacuole, peroxisome, thylakoid, stroma, granum, nuclear envelope, nucleolus, flagellum, cytoskeleton, microvilli, chromatin, plasmodesmata, microfilaments, microtubules, fimbriae, nucleoid, capsule, flagella, nucleoid, plasma membrane, cytosol, phagocytosis, endocytosis, cristae.</td></tr>
<tr><td colspan="4">Phenomena</td></tr>
<tr><td>Context/ Phenomena</td><td colspan="3">Some example phenomena for MS-LS1-2:
• Skin cells act as a barrier between your insides and the outside.
• Under a microscope, a muscle cell looks different than a skin cell.
• Under a microscope, a root cell looks different than a leaf cell.
• An E. coli bacterium is approximately the same size as the mitochondria of a mammalian lung cell.</td></tr>
<tr><td colspan="4">This Performance Expectation and associated Evidence Statements support the following Task Demands.</td></tr>
<tr><td colspan="4">Task Demands</td></tr>
<tr><td colspan="4">1. Assemble or complete, from a collection of potential model components, an illustration that is capable of representing a eukaryotic (plant and/or animal) or prokaryotic cell in terms of the function of the cell.</td></tr>
<tr><td colspan="4">2. Select or identify from a collection of potential model components, including distractors, the components needed to model the phenomenon. Components might mirror the cell wall, cell membrane, nucleus, chloroplast, and/or mitochondrion. This <u>does not</u> include labeling an existing diagram.</td></tr>
</table>

3.	Manipulate the components of a model to demonstrate the changes, properties, and/or events that act to result in the phenomenon.*
4.	Given models or diagrams of cells, identify the functions of each part of the cell.
5.	Identify missing components, relationships, or other limitations of the model.
6.	Describe, select, or identify the relationships among components of a model that together function as a cell.

*denotes those task demands which are deemed appropriate for use in stand-alone item development

<table>
<tr><td>Performance Expectation</td><td colspan="3">MS-LS1-3
Use argument supported by evidence for how the body is a system of interacting sub-systems composed of groups of cells.</td></tr>
<tr><td>Dimensions</td><td>Engaging in Argument from Evidence
• Use an oral and written argument supported by evidence to support or refute an explanation or a model for a phenomenon.</td><td>LS1.A: Structure and Function
• In multicellular organisms, the body is a system of multiple interacting subsystems. These subsystems are groups of cells that work together to form tissues and organs that are specialized for particular body functions.</td><td>Systems and System Models
• Systems may interact with other systems; they may have sub-systems and be part of larger complex systems.</td></tr>
<tr><td>Clarifications and Content Limits</td><td colspan="3">Clarification Statements
• Emphasis is on the conceptual understanding that cells form tissues and tissues form organs specialized for particular body functions. Examples could include the interaction of subsystems within a system and the normal functioning of those systems.

Content Limits
• Assessment does not include the mechanism of one body system independent of others.
• Assessment is limited to the circulatory, excretory, digestive, respiratory, muscular and nervous systems.</td></tr>
<tr><td>Science Vocabulary Students are Expected to Know</td><td colspan="3">Organization, organ, organ system, response, internal cue, life-sustaining functions, muscular system, anatomy, aorta, artery, automatic, bone, bone marrow, brain, brain stem, cerebellum, cerebrum, circulatory system, connective tissue, cornea, digestive system, gland, lens, muscle, muscle cell, reflex, sensory, skeletal system, tissue, respiratory, vertebrate, invertebrate, reproduction, breed, heart, lungs, heart rate</td></tr>
<tr><td>Science Vocabulary Students are Not Expected to Know</td><td colspan="3">Destabilize, excitatory molecule, feedback mechanism, hierarchical, homeostasis, inhibitory molecule, immune system, living system, neural, organic compound synthesis, protein structure, protein synthesis, regulate, stabilize, stomate, system level, transform/transport matter/and or energy, excretion, limiting factor, voluntary muscle, pancreas, sensory fiber, sensory nerve, root development</td></tr>
<tr><td colspan="4">Phenomena</td></tr>
<tr><td>Context/ Phenomena</td><td colspan="3">Some example phenomena for MS-LS1-3:
• After falling and scraping your knee, a scab forms over the wound.
• An elephant’s heart rate is slower than a mouse’s heart rate even though it is much bigger.
• A person swallows their food while doing a handstand, but a bird cannot swallow food while hanging upside down.
• When a person hasn’t eaten in a few hours and is hungry, their stomach makes an audible “growling” sound.</td></tr>
<tr><td colspan="4">This Performance Expectation and associated Evidence Statements support the following Task Demands.</td></tr>
<tr><td colspan="4">Task Demands</td></tr>
<tr><td colspan="4">1. Based on the provided data, identify, describe or illustrate a claim regarding the relationship between cells, tissues, organs and bodily function(s).</td></tr>
<tr><td colspan="4">2. Identify, summarize, or organize given data or other information to support or refute a claim regarding the relationship between cells, tissues, organs and bodily function(s).*</td></tr>
</table>

3. Sort inferences about the relationship between body systems into those that are supported by the data, contradicted by the data, or neither, or some similar classification.*
4. Select supporting evidence from competing sources based on the reliability of statistical relationships, how representative the sample is, or study design to show how the body is a system of interacting subsystems.
5. Construct an argument using scientific reasoning drawing on credible evidence to explain the relationships of interacting subsystems in a body such as tissues and organs. (Hand scored CR) *
6. Identify additional evidence that would help clarify, support, or contradict a hypothesized relationship or causal argument regarding the interactions of subsystems in the body.
7. Identify or describe alternate explanations and the data needed to distinguish among them in order to explain how body system functions.

*denotes those task demands which are deemed appropriate for use in stand-alone item development

<table>
<tr><td>Performance Expectation</td><td colspan="3">MS-LS1-4
Use argument based on empirical evidence and scientific reasoning to support an explanation for how characteristic animal behaviors and specialized plant structures affect the probability of successful reproduction of animals and plants respectively.</td></tr>
<tr><td>Dimensions</td><td>Engaging in Argument from Evidence
• Use an oral and written argument supported by empirical evidence and scientific reasoning to support or refute an explanation or a model for a phenomenon or a solution to a problem.</td><td>LS1.B: Growth and Development of Organisms
• Animals engage in characteristic behaviors that increase the odds of reproduction.
• Plants reproduce in a variety of ways, sometimes depending on animal behavior and specialized features for reproduction.</td><td>Cause and Effect
• Phenomena may have more than one cause, and some cause and effect relationships in systems can only be described using probability.</td></tr>
<tr><td>Clarifications and Content Limits</td><td colspan="3">Clarification Statements:
• Examples of behaviors that affect the probability of animal reproduction could include: nest building to protect young from cold, herding of animals to protect young from predators, and vocalization of animals and colorful plumage to attract mates for breeding.
• Examples of animal behaviors that affect the probability of plant reproduction could include: transferring pollen or seeds, and creating conditions for seed germination and growth.
• Examples of plant structures could include: bright flowers attracting butterflies that transfer pollen, flower nectar and odors that attract insects that transfer pollen, and hard shells on nuts that squirrels bury.

Content Limits:
• Data analysis should be limited to calculations and interpretation of measures of central tendency.
• Students are only expected to understand probability as expected relative frequency.
• Students can be asked to evaluate whether sample data are representative and the limits to which findings can be generalized.
• Data sets can include not only common trends but also outliers and anomalous data points.
• <u>Students do not need to know</u>: Mechanisms or patterns of inheritance, meiosis, specific reproductive structures not detailed within this document (e.g., nuptial pads, dulap), detailed life cycles.</td></tr>
<tr><td>Science Vocabulary Students are Expected to Know</td><td colspan="3">Nest, herd, mate, breed, probability, behavior, pollen, flower, petal, seed, fruit, nectar, germination, vocalization, plumage, pollination</td></tr>
<tr><td>Science Vocabulary Students are Not Expected to Know</td><td colspan="3">Symbiosis, mutualism, commensalism, parasitism, gametophyte, sporophyte, carpal, sepal, pistil, anther, stamen, ovule, “alternation of generations,” sporangia, monoecious, dioecious.</td></tr>
<tr><td colspan="4" align="center">Phenomena</td></tr>
<tr><td>Context/ Phenomena</td><td colspan="3">Some example phenomena for MS-LS1-4:
• Spring peepers (Pseudacris crucifer) in South Georgia, North Georgia, and Eastern Kentucky begin vocalizing (breeding) at different times of the year.</td></tr>
</table>

	• Female poison arrow frogs lay their eggs in leaf litter. When they hatch, male poison arrow frogs herd the tadpoles onto their backs and transport them to bromeliads, where they develop into adulthood. • The proportion of trees that are pollinated by insects decreases with latitude (phenomenon would be data tables that illustrate this relationship). • The Aspen tend to be one of the first plants to emerge after a forest fire.
This Performance Expectation and associated Evidence Statements support the following Task Demands.	
Task Demands	
1. Based on the provided data, identify, describe or illustrate a claim regarding the relationship between a characteristic animal behavior and/or specialized plant structure and the probability of successful reproduction in the species.	
2. Identify, summarize, or organize given data or other information to support or refute a claim regarding the relationship between a characteristic animal behavior and/or specialized plant structure and the probability of successful reproduction in the species.	
3. Sort inferences about the relationship of behaviors or structures to breeding success into those that are supported by the data, contradicted by the data, or neither, or some similar classification.	
4. Select supporting evidence from competing sources based on the reliability of statistical relationships, how representative the sample is, or study design.	
5. Construct an argument using scientific reasoning drawing on credible evidence to explain the relationships of animal behaviors or plant structures to reproductive success. (Hand scored CR)	
6. Identify additional evidence that would help clarify, support, or contradict a hypothesized relationship or causal argument.	
7. Identify or describe alternate explanations and the data needed to distinguish among them.	

<table>
<tr><td>Performance Expectation</td><td colspan="3">MS LS1-5
Construct a scientific explanation based on evidence for how environmental and genetic factors influence the growth of organisms.</td></tr>
<tr><td>Dimensions</td><td>Constructing Explanations and Designing Solutions
• Construct a scientific explanation based on valid and reliable evidence obtained from sources (including the students’ own experiments) and the assumption that theories and laws that describe the natural world operate today as they did in the past and will continue to do so in the future.</td><td>LS1.B: Growth and Development of Organisms
• Genetic factors as well as local conditions affect the growth of the adult plant.</td><td>Cause and Effect
• Phenomena may have more than one cause, and some cause and effect relationships in systems can only be described using probability.</td></tr>
<tr><td>Clarifications and Content Limits</td><td colspan="3">Clarification Statements
• Examples of local environmental conditions could include availability of food, light, space, and water.
• Examples of genetic factors could include large breed cattle and species of grass affecting growth of organisms.
• Examples of evidence could include a drought decreasing plant growth, fertilizer increasing plant growth, different varieties of plant seeds growing at different rates in different conditions, and fish growing larger in large ponds than they do in small ponds.

Content Limits
• Assessment does not include genetic mechanisms, gene regulation, or biochemical processes.
• Assessment does not include Punnett squares.
• <u>Students do not need to know</u>: epigenetics or variations of gene expression.</td></tr>
<tr><td>Science Vocabulary Students are Expected to Know</td><td colspan="3">Gene, genetics, genome, genotype, phenotype, environment, growth, development, DNA.</td></tr>
<tr><td>Science Vocabulary Students are Not Expected to Know</td><td colspan="3">Epigenetics, RNA, gene expression, photoperiod.</td></tr>
<tr><td colspan="4">Phenomena</td></tr>
<tr><td>Context/ Phenomena</td><td colspan="3">Phenomena for this performance expectation should include two groups of a particular organism with one environmental change.

Some example phenomena for MS-LS1-5:
• An orchard contains both full-sized and dwarf apple trees. Individuals of both types of tree grow shorter and produce fewer apples when planted on a dry hillside, and grow taller and produce more apples when planted on the shore of a pond. (i.e., the full apple trees on the hillside are the same size with similar apple production as the dwarf apple trees by the pond).
• Only about 90% of identical twins each have the same height.</td></tr>
</table>

	• A group of poinsettias and daisies are grown in the same greenhouse. The poinsettias bloom when exposed to ten consecutive hours of light, but the daisies bloom when exposed to 14 consecutive hours of light. • Burrs are dispersed to different environments by traveling on the fur of mammals. Some seeds from a burr plant drop off into a sunny field, while others drop off into a shady patch of woods. The burr plants that grew in the sun are taller and produced more burrs than those that grew in the shade.

This Performance Expectation and associated Evidence Statements support the following Task Demands.
Task Demands
1. Articulate, describe, illustrate, or select genetic and/or environmental influences on phenotypic differences between organisms. This may entail sorting relevant from irrelevant information.
2. Explain the process by which genetic factors and/or local conditions cause the observed phenomenon, supporting the explanation with valid and reliable evidence (hand scored).
3. Identify evidence that supports the inference that genetic and environmental factors influence growth and development of organisms. Environmental factors may include food, light, space, and water.
4. Describe, identify, and/or select information from one or more sources to support an explanation for phenotypic differences in organisms related to genetic and environmental factors.

<table>
<tr><td>Performance Expectation</td><td colspan="3">MS-LS1-6
Construct a scientific explanation based on evidence for the role of photosynthesis in the cycling of matter and flow of energy into and out of organisms.</td></tr>
<tr><td>Dimensions</td><td>Constructing Explanations and Designing Solutions
• Construct a scientific explanation based on valid and reliable evidence obtained from sources (including the students' own experiments) and the assumption that theories and laws that describe the natural world operate today as they did in the past and will continue to do so in the future.</td><td>LS1.C: Organization for Matter and Energy Flow in Organisms
• Plants, algae (including phytoplankton), and many microorganisms use the energy from light to make sugars (food) from carbon dioxide from the atmosphere and water through the process of photosynthesis, which also releases oxygen. These sugars can be used immediately or stored for growth or later use.

PS3.D: Energy in Chemical Processes and Everyday Life
• The chemical reaction by which plants produce complex food molecules (sugars) requires an energy input (i.e., from sunlight) to occur. In this reaction, carbon dioxide and water combine to form carbon-based organic molecules and release oxygen (secondary).</td><td>Energy and Matter
• Within a natural system, the transfer of energy drives the motion and/or cycling of matter.</td></tr>
<tr><td>Clarifications and Content Limits</td><td colspan="3">Clarification Statements
• Emphasis is on tracing movement of matter and flow of energy.
• Students are able to identify relationships between dependent and independent variables.

Content Limits
• Assessment does not include the biochemical mechanisms of photosynthesis.
• Assessment does not include the carbon cycle or nitrogen fixation.
• <u>Students do not need to know:</u> how to balance chemical equations.</td></tr>
<tr><td>Science Vocabulary Students are Expected to Know</td><td colspan="3">Glucose, algae, consumer, product, transformation, conservation, convert, decomposer, aquatic, organic, phytoplankton, producer, reaction, carbon, carbon dioxide, chemical process, chemical reaction, molecule, nutrient, moisture, structure, organic matter, stimulus, tissue, hydrogen</td></tr>
<tr><td>Science Vocabulary Students are Not Expected to Know</td><td colspan="3">Biomass, biological molecule, compound, flow of matter, hydrocarbon, net transfer, photosynthesizing organism, carbon cycle, efficient, excitatory molecule, molecular synthesis, organic compound synthesis, stomata</td></tr>
<tr><td colspan="4">Phenomena</td></tr>
<tr><td>Context/ Phenomena</td><td colspan="3">Some example phenomena for MS-LS1-6:
• A plant is kept in a clear, closed container that allows sunlight to pass through. After one week, the plant is dead. A mouse kept alone in the same container also dies. However, a plant and mouse kept together in the same container after one week are alive.
• The plant Elodea releases bubbles at an increased rate when an aquatic animal is added to the same aquarium.
• A plant grows in a pot of soil for one month. Only water is added to the pot. After one month, the plant has gained mass, while the mass of the soil has barely changed.</td></tr>
</table>

	• A plant leaf kept in the light contains large amounts of starch, while a leaf kept in the dark does not.
This Performance Expectation and associated Evidence Statements support the following Task Demands.	
Task Demands	
1. Articulate, describe, illustrate, or select the relationships, interactions, and/or processes to be explained. This may entail sorting relevant from irrelevant information or features of the reactants and products.	
2. Express or complete a description of the flow of energy and/or matter among organisms. This may include indicating directions of causality in an incomplete model (including food webs), such as a flow chart or diagram.	
3. Identify evidence that photosynthesis cycles matter and energy through an ecosystem.	
4. Select, identify, or describe the predicted effect of a change of conditions on the flow of energy and matter among organisms.	
5. Describe, identify, and/or select information needed to support an explanation.	

<table>
<tr><td>Performance Expectation</td><td colspan="3">MS-LS1-7
Develop a model to describe how food is rearranged through chemical reactions to form new molecules that support growth and/or release energy as this matter moves through an organism.</td></tr>
<tr><td>Dimensions</td><td>Developing and Using Models
• Develop a model to describe unobservable mechanisms.</td><td>LS1.C: Organization for Matter and Energy Flow in Organisms
• Within individual organisms, food moves through a series of chemical reactions in which it is broken down and rearranged to form new molecules, support growth, or release energy.

PS3.D: Energy in Chemical Processes and Everyday Life
• Cellular respiration in plants and animals involves chemical reactions with oxygen that release stored energy. In these processes, complex molecules containing carbon react with oxygen to produce carbon dioxide and other materials. (secondary)</td><td>Energy and Matter
• Matter is conserved because atoms are conserved in physical and chemical processes.</td></tr>
<tr><td>Clarifications and Content Limits</td><td colspan="3">Clarification Statements
• Emphasis is on the describing that describing that molecules are broken apart and put back together and that in this process energy is released.

Content Limits
• Assessment does not include details of the chemical reactions for photosynthesis or respiration.
• <u>Students do not need to know</u>: enzymes, ATP synthase, metabolism, biochemical pathways, redox reactions, molecular transport, specific enzymes involved, catalysts</td></tr>
<tr><td>Science Vocabulary Students are Expected to Know</td><td colspan="3">Oxygen, carbon dioxide, water, sugar, glucose (including chemical formula), ATP, chemical bonds, photosynthesis, proteins, enzymes, organelles, nucleus, DNA, mitochondria, cytosol, cytoplasm, nitrogen</td></tr>
<tr><td>Science Vocabulary Students are Not Expected to Know</td><td colspan="3">Biochemical, fatty acids, oxidizing agent, electron acceptor, biosynthesis, locomotion, phosphorylation, electron transport chain, chemiosmosis, pyruvate, pentose, adenine, phosphate, amino acid, fermentation, aerobic respiration, redox reactions, oxidation, reduction, reducing agent, oxidizing agent, NAD+, transport chain, glycolysis, citric acid cycle, oxidative phosphorylation, substrate-level phosphorylation, acetyl CoA, cytochromes, chemiosmosis, ATP synthase, lactic acid</td></tr>
<tr><td colspan="4">Phenomena</td></tr>
<tr><td>Context/ Phenomena</td><td colspan="3">Some example phenomena for MS-LS1-7:
• A young plant is grown in a bowl of sugar water. As it grows, the amount of sugar in the water decreases.
• A person feels tired and weak before they eat lunch. After they eat some fruit, they feel more energetic and awake.
• An athlete completing difficult training feels that their muscles recover and repair faster when they eat more high-protein foods in a day compared to when they eat less protein in a day.
• Amoeba are provided food in a petri dish. When fed, the amoeba become more active and begin to grow and divide</td></tr>
<tr><td colspan="4">This Performance Expectation and associated Evidence Statements support the following Task Demands.</td></tr>
<tr><td colspan="4">Task Demands</td></tr>
</table>

1. Select or identify from a collection of potential model components, including distractors, the components needed to model the phenomenon. Components might include gases, sugars, and organelles.
2. Assemble or complete, from a collection of potential model components, an illustration or flow chart that is capable of representing the transformation of food + oxygen into energy and/or new compounds. This does not include labeling an existing diagram.
3. Manipulate the components of the model to demonstrate the changes, properties, processes, and/or events that act to result in the phenomenon.
4. Make predictions about the effects of changes in the type or amount of a certain component in the model. Predictions can be made by manipulating model components, completing illustrations, or selecting from lists with distractors.
5. Given models or diagrams of the state of model components, identify the properties of the system that give rise to the phenomenon.
6. Identify missing components, relationships, or other limitations of the model.
7. Describe, select, or identify the relationships among components of a model that describe or explain how food can be turned into energy for new growth and other activities.

<table>
<tr><td>Performance Expectation</td><td colspan="3">MS-LS1-8
Gather and synthesize information that sensory receptors respond to stimuli by sending messages to the brain for immediate behavior or storage as memories.</td></tr>
<tr><td>Dimensions</td><td>Obtaining, Evaluating, and Communicating Information
• Gather, read, and synthesize information from multiple appropriate sources and assess the credibility, accuracy, and possible bias of each publication and methods used, and describe how they are supported or not supported by evidence.</td><td>LS1.D: Information Processing
• Each sense receptor responds to different inputs (electromagnetic, mechanical, chemical), transmitting them as signals that travel along nerve cells to the brain. The signals are then processed in the brain, resulting in immediate behaviors or memories.</td><td>Cause and Effect
• Cause and effect relationships may be used to predict phenomena in natural systems.</td></tr>
<tr><td>Clarifications and Content Limits</td><td colspan="3">Content Limits
• Assessment does not include mechanisms for the transmission of information from sensory receptors to the brain.
• <u>Students do not need to know</u>: Sensory transduction, ion channels, action potentials, sensory and motor cortices in the brain.</td></tr>
<tr><td>Science Vocabulary Students are Expected to Know</td><td colspan="3">Memory, perception, transmit, accuracy, immediate, nerve, receptor, sense receptor, sensory, behavioral response to stimuli, electromagnetic, stimulus, short-term memory, long-term memory, salt, sour, sweet, bitter, brain, nervous system, taste, smell, touch, hear, sight</td></tr>
<tr><td>Science Vocabulary Students are Not Expected to Know</td><td colspan="3">Neuron, neurotransmitter, endocrine signaling, synapse, axon, olfactory, rods, cones, trichromatic vision, retina, hair cells, cochlea, fight-or-flight response, sensitization, depolarization, taste papillae, umami,</td></tr>
<tr><td colspan="4">Phenomena</td></tr>
<tr><td>Context/ Phenomena</td><td colspan="3">Some example phenomena for MS-LS1-8:
• A woman closes her eyes and touches the tip of her nose with her index finger.
• A student is studying in a library. The fire alarm goes off and he involuntarily jumps out of his chair.
• A woman walking past a bakery smells cinnamon and is instantly reminded of her grandmother’s house.
• A driver sees a stoplight change from green to red and quickly moves his foot from the accelerator pedal to the break.</td></tr>
<tr><td colspan="4">This Performance Expectation and associated Evidence Statements support the following Task Demands.</td></tr>
<tr><td colspan="4">Task Demands</td></tr>
<tr><td colspan="4">1. Analyze and interpret scientific evidence from multiple scientific/technical sources including text, diagrams, charts, symbols and mathematical representations to describe how external stimuli are sensed by the brain.</td></tr>
<tr><td colspan="4">2. Assemble or complete an illustration or flow chart representing physiological or behavioral responses to external stimuli.</td></tr>
<tr><td colspan="4">3. Based on the information provided, identify or describe supporting evidence for an argument regarding the relationship between an external stimulus, sensory receptors and/or a particular behavior.</td></tr>
</table>

4. Make predictions about the effects on sensory receptors, immediate behavior, or memory storage as a result of changes to an external stimulus. Predictions can be quantitative or qualitative and can be made by completing illustrations, or selecting from lists with distractors.
5. Evaluate the validity, credibility, accuracy, relevancy and/or possible bias of scientific/technical sources.
6. Synthesize an explanation regarding sensory stimuli that incorporates scientific evidence from multiple sources.
7. Identify, summarize, or organize given data or other information to support or refute a claim relating the characteristics of an external stimulus to a sensory pathway.

<table>
<tr><td>Performance Expectation</td><td colspan="3">MS-LS2-1
Analyze and interpret data to provide evidence for the effects of resource availability on organisms and populations of organisms in an ecosystem.</td></tr>
<tr><td>Dimensions</td><td>Analyzing and Interpreting Data
• Analyze and interpret data to provide evidence of phenomena.</td><td>LS2.A: Interdependent Relationships in Ecosystems
• Organisms, and populations of organisms, are dependent on their environmental interactions, both with other living things and with nonliving factors.
• In any ecosystem, organisms and populations with similar requirements for food, water, oxygen, or other resources may compete with each other for limited resources, access to which consequently constrains their growth and reproduction.
• Growth of organisms and population increases are limited by access to resources.</td><td>Cause and Effect
• Cause and effect relationships may be used to predict phenomena in natural or designed systems.</td></tr>
<tr><td>Clarifications and Content Limits</td><td colspan="3">Clarification Statements
• Emphasis is on cause and effect relationships between resources and growth of individual organisms, and the numbers of organisms in ecosystems during periods of abundant and scarce resources.
• Examples could include water, food, and living space

Content Limits
• Assessment does not include mathematical and/or computational representations of factors related to carrying capacity of ecosystems of different sizes (including deriving mathematical equations to make comparisons).</td></tr>
<tr><td>Science Vocabulary Students are Expected to Know</td><td colspan="3">Resource, competition, ecosystem, nutrient, food chain/web, producer, consumer</td></tr>
<tr><td>Science Vocabulary Students are Not Expected to Know</td><td colspan="3">Biotic component, abiotic component, exponential (AKA “logistic”) growth, ecological niche, resource partitioning, fundamental niche, realized niche, carrying capacity, interspecific competition, intraspecific competition, biomass, carrying capacity</td></tr>
<tr><td colspan="4">Phenomena</td></tr>
<tr><td>Context/ Phenomena</td><td colspan="3">The phenomena for this performance expectation are the given data. Samples of phenomena should describe the data set(s) to be given in terms of patterns or relationships to be found in the data, and the columns and rows of a hypothetical table presenting the data, even if the presentation is not tabular. The description of the phenomenon should describe the presentation format of the data (e.g., maps, tables, graphs, etc.).

Some example phenomena for MS-LS2-1:
• On the north Atlantic coastline, two species of barnacles live at different depths
• Cheetahs and leopards in the savannah use the same watering holes.
• After a drought period, the population of grasshoppers is halved.
• A garden is cleared of aphids. After a few days, the ladybirds in the surrounding trees are gone.</td></tr>
</table>

This Performance Expectation and associated Evidence Statements support the following Task Demands.
Task Demands
6. Organize and/or arrange (e.g., using illustrations and/or labels), or summarize data to highlight trends, patterns, or correlations between resource availability and the growth of a population or populations of organisms.
7. Generate or construct graphs, tables, or assemblages of illustrations and/or labels of data that document patterns, trends, or correlations between resource availability and the growth of a population or populations of organisms. This may include sorting out distractors.*
8. Use relationships identified in resource/population data to predict the change in a population or populations or the change in resources that resulted in a change in populations.**
9. Identify patterns or evidence in the data that supports inferences and explanations about how resource availability affects a population of organisms.*
10. Construct or identify testable questions that can be asked to collect data about how resource availability may affect the growth of a population or populations of organisms.
11. Identify, describe, or select from a collection characteristics to be manipulated or held constant while gathering information to answer a well-articulated question.*
12. Select or describe inferences relevant to the question posed and supported by the data, especially inferences about causes and effects.
13. Select, identify, or describe predicted outcomes when specific changes in resource availability occur, using inferences about cause and effect relationships involving those resources.**

*denotes those task demands which are deemed appropriate for use in stand-alone item development
**TD3 and TD8 must be used together.

Performance Expectation	**MS-LS2-2** Construct an explanation that predicts patterns of interactions among organisms across multiple ecosystems.		
Dimensions	**Constructing Explanations and Designing Solutions** • Construct an explanation that includes qualitative or quantitative relationships between variables that predict phenomena.	**LS2.A: Interdependent Relationships in Ecosystems** • Similarly, predatory interactions may reduce the number of organisms or eliminate whole populations of organisms. Mutually beneficial interactions, in contrast, may become so interdependent that each organism requires the other for survival. Although the species involved in these competitive, predatory, and mutually beneficial interactions vary across ecosystems, the patterns of interactions of organisms with their environments, both living and nonliving, are shared.	**Patterns** • Patterns can be used to identify cause and effect relationships.
Clarifications and Content Limits	**Clarification Statement** • Emphasis is on predicting consistent patterns of interactions in different ecosystems in terms of the relationships among and between living organisms and nonliving components of ecosystems. • Examples of types of interactions could include competitive, predatory, and mutually beneficial. **Content Limits** • Analysis may include recognizing patterns in data, specifying and explaining relationships, making logical predictions from data, retrieving information from a table, graph or figure and using it to explain relationships, generating hypotheses based on observations or data, and generalizing a pattern. • Analysis should not include relating mathematical or scientific concepts to other content areas.		
Science Vocabulary Students are Expected to Know	relative, disperse, ecological role, host, infection, mutualism, mutually beneficial, parasite, evolve, genetic, interdependent		
Science Vocabulary Students are Not Expected to Know	abiotic		
Phenomena			
Context/ Phenomena	For this performance expectation, the phenomena are sets of data. Those are the observed facts that the students will look at to discover patterns. Below, we enumerate some of the patterns that might comprise the data sets (phenomena) to be analyzed. Patterns should be observed across at least **two** different environments/habitats. Patterns that describe the data sets for MS-LS2-2: • The tongue of the alligator snapping turtle looks like a small worm. The turtle uses this tongue to lure prey close to its mouth. (Predation)—also angler fish.		

	• Higher density of squirrels in oak environment than in maple environment. • Hippopotamuses spend time in both aquatic and savannah ecosystems. When found in aquatic environments, they're often surrounded by carp. When found in a savannah environment, they're often surrounded by oxpeckers. • In Ecuador's Andean Cloud Forest, a hummingbird feeds on the nectar of an orchid flower (*Epidendrum secundum*). In the Madagascar, a similar orchid flower (*Angraecum sesquipedale*) is seen, but no hummingbirds are found.
This Performance Expectation and associated Evidence Statements support the following Task Demands.	
Task Demands	
1. Articulate, describe, illustrate, or select the relationships or interactions to be explained. This may entail sorting relevant from irrelevant information or features.	
2. Express or complete a causal chain common or distinct across organisms or environments. This may include indicating directions of causality in an incomplete model such as a flow chart or diagram or completing cause and effect chains.*	
3. Identify evidence supporting the inference of causation of patterns of interactions among organisms across multiple ecosystems expressed in a causal chain.*	
4. Use an explanation to predict interactions among different organisms or in different environments.	
5. Describe/Identify/Select information needed to support an explanation of patterns of interactions among organisms across multiple ecosystems.	

*denotes those task demands which are deemed appropriate for use in stand-alone item development

Performance Expectation	**MS-LS2-3** Develop a model to describe the cycling of matter and flow of energy among living and nonliving parts of an ecosystem.		
Dimensions	**Developing and Using Models** • Develop a model to describe phenomena.	**LS2.B: Cycle of Matter and Energy Transfer in Ecosystems** • Food webs are models that demonstrate how matter and energy are transferred among producers, consumers, and decomposers as the three groups interact within an ecosystem. Transfers of matter into and out of the physical environment occur at every level. Decomposers recycle nutrients from dead plant or animal matter back to the soil in terrestrial environments or to the water in aquatic environments. The atoms that make up the organisms in an ecosystem are cycled repeatedly between the living and nonliving parts of the ecosystem.	**Energy and Matter** • The transfer of energy can be tracked as energy flows through a natural system.

Clarifications and Content Limits	**Clarification Statements**
	• Emphasize food webs and the role of producers, consumers, and decomposers in various ecosystems.
	• Emphasis is on describing the conservation of matter and flow of energy into and out of various ecosystems, and on defining the boundaries of the system.
	
	Content Limits
	• Students do not need to identify biomes or to know information about specific biomes.
	• Assessment does not include
	  o The use of chemical reactions to describe the processes.
	  o Identification of trophic levels, understanding of the relative energies of the trophic levels, nor the knowledge of the 10% energy transfer between trophic levels.
	  o The process of bioaccumulation.		

Science Vocabulary Students Are Expected to Know	Producer, consumer, decomposer, herbivore, omnivore, carnivore, algae, fungi, microbe, microorganism, organic matter, organic waste, photosynthesis, atom, molecule, sugar, carbon, carbon dioxide, nitrogen, oxygen, predator, prey, aquatic, interdependent, chemical reaction, reactant, product		
Science Vocabulary Students Are Not Expected to Know	Biotic, abiotic, trophic level, energy pyramid, nitrogen fixation, exothermic/endothermic, detritivores, biomass, bioaccumulation/biomagnification, autotroph/heterotroph, biosphere, hydrosphere, geosphere, aerobic, anaerobic, phosphorous, phytoplankton.		
Phenomena			
Context/ Phenomena	Some example phenomena for MS-LS2-3: • In the Alaskan tundra, more grass and wildflowers grow on top of underground fox dens than elsewhere. • In July, a colony of lava crickets is found to inhabit lava flows from a May eruption, but the first plant does not appear in the area until November. • Fox-inhabited islands in the Aleutian Islands have less vegetation than islands not inhabited by foxes. • Giant clams and tube worms are found in the darkest parts of the oceans in the hot water near hydrothermal vents.		

This Performance Expectation and associated Evidence Statements support the following Task Demands.
Task Demands
1. Identify, assemble, or complete from a collection of potential model components, including distractors, components of a food-web model that describe transfers of matter and/or energy among producers, consumers, decomposers, or some subsets of those, potentially including transfers between living and nonliving organisms.
2. Describe, select, or identify the relationships among components of a food-web model that describes how parts of the food web (producers, consumers, and decomposers) interact to continually cycle matter and to transfer energy among living and nonliving parts of an ecosystem.
3. Manipulate the components of a food-web model to demonstrate how the interactions among producers, consumers, and/or decomposers result in changes to the cycling of matter and/or transfer of energy among living and nonliving parts of an ecosystem.
4. Select, describe, or illustrate predictions about the effects of changes in the organisms or nonliving components of the environment on the cycling of matter, transfer of energy, and/or other organisms in the environment. Predictions can be made by manipulating model components, completing illustrations, or selecting from lists with distractors.
5. Select or identify missing components or relationships of a food web model that describes the transfers of matter and/or energy among living and nonliving parts of an ecosystem.

<table>
<tr><td>Performance Expectation</td><td colspan="3">MS-LS2-4
Construct an argument supported by empirical evidence that changes to physical or biological components of an ecosystem affect populations.</td></tr>
<tr><td>Dimensions</td><td>Engaging in Argument from Evidence
• Construct an oral and written argument supported by empirical evidence and scientific reasoning to support or refute an explanation or a model for a phenomenon or a solution to a problem.</td><td>LS2.C: Ecosystem Dynamics, Functioning, and Resilience
• Ecosystems are dynamic in nature: their characteristics can vary over time. Disruptions to any physical or biological component of an ecosystem can lead to shifts in all its populations.</td><td>Stability and Change
• Small changes in one part of a system might cause large changes in another part.</td></tr>
<tr><td>Clarifications and Content Limits</td><td colspan="3">Clarification Statements
• Emphasis is on recognizing patterns in data and making warranted inferences about changes in populations, and on evaluating empirical evidence supporting arguments about changes to ecosystems.

Content Limits
• Assessment does not include the use of chemical reactions to describe the processes.</td></tr>
<tr><td>Science Vocabulary Students are Expected to Know</td><td colspan="3">Predator, prey, mutually beneficial interactions, competition, consumers, producers, decomposers, biodiversity.</td></tr>
<tr><td>Science Vocabulary Students are Not Expected to Know</td><td colspan="3">Carrying capacities, anthropogenic changes, biomass</td></tr>
<tr><td colspan="4">Phenomena</td></tr>
<tr><td>Context/ Phenomena</td><td colspan="3">Example Phenomena for MS-LS2-4:
• After a beaver builds a dam, the amount and diversity of fish life in a stream increases.
• After wolves were reintroduced to Yellowstone, there were more willows.
• The number of willows has increased in Yellowstone. (Give two competing hypotheses: wolf introduction; beaver population increase).
• As the Aral Sea declined in size since the 1960s, salinity has increased and the Aral trout is no longer present in the lake.</td></tr>
<tr><td colspan="4">This Performance Expectation and associated Evidence Statements support the following Task Demands.</td></tr>
<tr><td colspan="4">Task Demands</td></tr>
<tr><td colspan="4">1. Articulate, describe, illustrate, or select the relationships, interactions, and/or processes to be explained. This may entail sorting relevant from irrelevant information or information supporting/refuting one or more competing hypotheses.</td></tr>
<tr><td colspan="4">2. Predict outcomes when changes to an ecosystem occur, given the inferred cause and effect relationships.*</td></tr>
<tr><td colspan="4">3. Identify, select, and/or describe information or evidence needed to support one or more potentially competing explanations.</td></tr>
</table>

4. Identify patterns of information/evidence in the data that support correlative/causative inferences about the relationships among the pertinent parts of an ecosystem.*
5. Organize and/or arrange (e.g., using illustrations and/or labels) or summarize population data to highlight trends, patterns, or correlations.

*denotes those task demands which are deemed appropriate for use in stand-alone item development

Performance Expectation	**MS-LS2-5** Evaluate competing design solutions for maintaining biodiversity and ecosystem services.		
Dimensions	**Engaging in Argument from Evidence** • Evaluate competing design solutions based on jointly developed and agreed-upon design criteria.	**LS2.C: Ecosystem Dynamics, Functioning, and Resilience** • Biodiversity describes the variety of species found in Earth's terrestrial and oceanic ecosystems. The completeness or integrity of an ecosystem's biodiversity is often used as a measure of its health. **LS4.D: Biodiversity and Humans** • Changes in biodiversity can influence humans' resources, such as food, energy, and medicines, as well as ecosystem services that humans rely on—for example, fresh air and water *(secondary).* **ETS1.B: Developing Possible Solutions** • There are systematic processes for evaluating solutions with respect to how well they meet the criteria and constraints of a problem *(secondary).*	**Stability and Change** • Small changes in one part of a system may cause a large change in another part.
Clarifications and Content Limits	**Clarification Statements** • Examples of ecosystem services could include water purification, nutrient recycling, and prevention of soil erosion. • Examples of design solution constraints could include scientific, economic, and social considerations. **Content Limits** • <u>Students do not need to know:</u> specific policies or specific details of organisms.		
Science Vocabulary Students Are Expected to Know	Habitats, niche, native species, non-native or invasive species		
Science Vocabulary Students Are Not Expected to Know	Specific species names, specific resource or habitat requirements for any species.		
Phenomena			
Context/ Phenomena	Engineering performance expectations are built around meaningful design problems rather than phenomena. In this case, the design problems involve preserving ecosystems and protecting biodiversity. For this performance expectation, the design problem and competing solutions replace phenomena. Some example design problems for MS-LS2-5: • Giant African Land Snails were brought to Florida by a boy who smuggled three snails into Florida. His grandmother released these into a garden and the snail population exploded. The snails eat over 500 plant species, tree bark, paint, and even stucco. Florida has implemented four solutions:		

	○ Trained dogs that sniff out snails for capture. ○ Chemicals applied to plants that the snails feed upon. ○ Predatory species to eat the snails. • The brown tree snake was accidentally brought to the island of Guam by ships during World War II, fed on native birds until the Guam rail, a native bird, nearly went extinct in 1984. Guam has implemented two solutions: ○ Feed rats acetaminophen and drop them into wooded areas. ○ Bring in predatory species to eat the snakes. • Cheatgrass, a type of weed that was brought to the United States in the late 1800s, has spread all over Utah from the desert valleys to the mountains, growing faster than most native plants. Utah has implemented two solutions: ○ Use genetically modified seeds for certain native seeds that are heartier than the Cheatgrass to push out the Cheatgrass seeds. ○ Controlled application of herbicides. • Asian carp is an aggressive fish species introduced in 1960 to control weed populations in waterways in southern fish farm ponds. The population was sterilized but a few fertile fish escaped into the Mississippi River and migrated north towards the Great Lakes. Asian carp are an invasive species that compete with native fish in the Great Lakes and threaten the ecosystem balance. Regions around the Great Lakes are implementing strategies: ○ Launch a campaign to encourage and incentivize fishing of Asian carp for human consumption ○ Use a system of electric barriers to prevent Asian carp from entering Lake Michigan from the Mississippi River. ○ Use nets to block paths to popular spawning sites during Asian carp reproduction season. ○ Introduce a botanic pesticide used for fish eradications in water areas known to have large Asian carp populations.

This Performance Expectation and associated Evidence Statements support the following Task Demands.
Task Demands
1. Identify or assemble from a collection, including distractors, the relevant aspects of the problem that, given design solutions if implemented, will resolve/improve maintaining biodiversity and ecosystem services.
2. Using given information for maintaining biodiversity and ecosystem services, select or identify constraints that the device or solution must meet.
3. Using the given information for maintaining biodiversity and ecosystem services, select or identify the criteria against which the device or solution should be judged.
4. Compare, rank, or otherwise evaluate the different design solutions for maintaining biodiversity and ecosystem services against the identified criteria.
5. Select or propose a recommended course of action supported by the design solution’s ability to meet identified criteria.

<table>
<tr><td>Performance Expectation</td><td colspan="3">MS-LS3-1
Develop and use a model to describe why structural changes to genes (mutations) located on chromosomes may affect proteins and may result in harmful, beneficial, or neutral effects to the structure and function of an organism.</td></tr>
<tr><td>Dimensions</td><td>Developing and Using a Model
• Develop and use a model to describe phenomena.</td><td>LS3.A: Inheritance of Traits
• Genes are located in the chromosomes of cells, with each chromosome pair containing two variants of each of many distinct genes. Each distinct gene chiefly controls the production of specific proteins, which in turn affects the traits of the individual. Changes (mutations) to genes can result in changes to proteins, which can affect the structures and functions of the organism and thereby change traits.

LS3.B: Variation of Traits
• In addition to variations that arise from sexual reproduction, genetic information can be altered because of mutations. Through rare, mutations my result in changes to the structure and function of proteins. Some changes are beneficial, others harmful, and some neutral to the organism.</td><td>Structure and Function
• Complex and microscopic structures and systems can be visualized, modeled, and used to describe how their function depends on the shapes, composition, and relationships among their parts; therefore, complex natural structures/systems can be analyzed to determine how they function.</td></tr>
<tr><td>Clarifications and Content Limits</td><td colspan="3">Clarification Statements
• Emphasis is on the conceptual understanding that changes in genetic material may result in making different proteins.

Content Limits
• Assessment does not include specific changes of genes at the molecular level, mechanisms for protein synthesis, and specific types of mutations.
• Do not use examples of mutations in humans.
• Analysis does not include species-level sources of genetic variation, including the founder effect, bottleneck, genetic drift or Hardy-Weinberg equilibrium.</td></tr>
<tr><td>Science Vocabulary Students are Expected to Know</td><td colspan="3">Genome, genotype, phenotype, DNA, pedigree, parent generation, trait, positive, negative, neutral, pollination, Punnett square, dominant trait, recessive trait, allele</td></tr>
<tr><td>Science Vocabulary Students are Not Expected to Know</td><td colspan="3">RNA, transcription, translation, mitosis, meiosis, interphase, prophase, metaphase, anaphase, telophase, cytokinesis, zygote, fertilization, codominance, incomplete dominance, sequencing, F1, F2, haploid, diploid, epigenetics, plasmid.</td></tr>
<tr><td colspan="4" align="center">Phenomena</td></tr>
<tr><td>Context/ Phenomena</td><td colspan="3">Some example phenomena for MS-LS3-1:
• Use of antibiotics in farming has leeched antibiotics into the water system. However, resistant bacteria persist in groundwater and are difficult to kill.</td></tr>
</table>

	• Wild almond trees produce the poisonous chemical amygdalin. Occasional individual almond trees have a mutation that cause them not to produce amygdalin. These individual plants are cultivated on almond farms. • A farmer observed one corn plant producing corn cobs with larger kernels. The farmer planted seeds from that plant and the offspring corn plants also had larger kernels. • Thale cress plants sprout in the spring and flower about a month later.

This Performance Expectation and associated Evidence Statements support the following Task Demands.
Task Demands
1. Select or identify from a collection of potential model components, including distractors, the components needed to model a phenomenon. Components might match a phenotypic change resulting from a mutation to various environments, to determine whether a mutation is beneficial, harmful, or neutral to the individual.
2. Assemble or complete, from a collection of potential model components, an illustration that is capable of representing the effects of a mutation in an individual in a specific environment. This <u>does not</u> include labeling an existing diagram.
3. Manipulate the components of a model to demonstrate the changes, properties, processes, and/or events that act to result in a phenomenon.
4. Make predictions about the effects of changes in an organism's ability to survive and reproduce based on the mutation and/or environment. Predictions can be made by manipulating model components, completing illustrations, or selecting from lists with distractors.
5. Given models or diagrams of phenotypic changes due to mutation, identify and describe why the mutation may positively, negatively, or neutrally affect the individual in different environments.
6. Identify or select the relationships among components of a model that describe the rationale behind the beneficial, harmful, or neutral nature of a mutation in specific environments.

<table>
<tr><td>Performance Expectation</td><td colspan="3">MS-LS3-2
Develop and use a model to describe why asexual reproduction results in offspring with identical genetic information and sexual reproduction results in offspring with genetic variation.</td></tr>
<tr><td>Dimensions</td><td>Developing and Using Models
• Develop and use a model to describe phenomena.</td><td>LS1.B: Growth and Development of Organisms
• Organisms reproduce, either sexually or asexually, and transfer their genetic information to their offspring (secondary).

LS3.A: Inheritance of Traits
• Variations of inherited traits between parent and offspring arise from genetic differences that result from the subset of chromosomes (and therefore genes) inherited.

LS3.B: Variation of Traits
• In sexually reproducing organisms, each parent contributes (at random) half of the genes acquired by the offspring. Individuals have two of each chromosome and hence two alleles of each gene, one acquired from each parent. These versions may be identical or may differ from each other.</td><td>Cause and Effect
• Cause and effect relationships may be used to predict phenomena in natural systems.</td></tr>
<tr><td>Clarifications and Content Limits</td><td colspan="3">Clarification Statements
• Emphasis is on using models such as Punnett Squares, diagrams and simulations to describe the cause and effect relationship of gene transmission from parent(s) to offspring and resulting genetic variation.

Content Limits
• Assessment does not include phases of mitosis or meiosis.
• <u>Students do not need to know</u>: process of recombination</td></tr>
<tr><td>Science Vocabulary Students are Expected to Know</td><td colspan="3">Development, germination, plant structure, plumage, reproductive system, fertilizer, allele, dominant trait, recessive trait, hereditary information, Punnett square, transmission, protein, DNA</td></tr>
<tr><td>Science Vocabulary Students are Not Expected to Know</td><td colspan="3">DNA replication, sex-linked trait, recombination, gene expression, segment, sex cell, sex chromosome, cell division, mutation, meiosis, amino acid, amino acid sequence, haploid, diploid</td></tr>
<tr><td colspan="4" align="center">Phenomena</td></tr>
<tr><td>Context/ Phenomena</td><td colspan="3">Some example phenomena for MS-LS3-2:
• Jellyfish will produce both clones and genetically diverse offspring during different stages of their life cycle.
• Strawberry plants grow another stem from a core stem that extends horizontally on the ground. This new stem will become a separate strawberry plant.
• A flatworm is cut in half. Rather than dying, both halves regenerate their lost portions to form two new, distinct, and fully functioning worms.</td></tr>
</table>

	• A plant (*Bryophyllum diagremontianum*) native to Madagascar has what appears to be miniature clusters of leaves lining the edges of a much larger leaf.

This Performance Expectation and associated Evidence Statements support the following Task Demands.
Task Demands
1. Select or identify from a collection of potential model components, including distractors, the components needed to model the phenomenon. Components might include alleles, genotypes, and phenotypes.
2. Assemble or complete, from a collection of potential model components, an illustration or flow chart that is capable of representing different types of reproduction. This *does not* include labeling an existing diagram.
3. Manipulate the components of a model to demonstrate the changes, properties, processes, and/or events that act to result in a phenomenon.
4. Make predictions about the effects of genetic variation from reproduction. Predictions can be made by manipulating model components, completing illustrations, or selecting from lists with distractors.
5. Given models or diagrams of types of reproduction, identify the types of reproduction and how they change in each scenario OR identify the properties of the different types of reproduction that cause genetic variation.
6. Identify missing components, relationships, or other limitations of the model.
7. Identify, calculate, or select the relationships among the components of a model that describe the types of reproduction, the environmental conditions under which reproduction occurs, or explain the genetic variation that results from reproduction.

<table>
<tr><td>Performance Expectation</td><td colspan="3">MS-LS4-1
Analyze and interpret data for patterns in the fossil record that document the existence, diversity, extinction, and change of life forms throughout the history of life on Earth, under the assumption that natural laws operate today as in the past.</td></tr>
<tr><td>Dimensions</td><td>Analyzing and Interpreting Data
• Analyze and interpret data to determine similarities and differences in findings.</td><td>LS4.A: Evidence of Common Ancestry and Diversity
• The collection of fossils and their placement in chronological order (e.g., through the location of the sedimentary layers in which they are found or through radioactive dating) is known as the fossil record. It documents the existence, diversity, extinction, and change of many life forms throughout the history of life on Earth.</td><td>Patterns
• Graphs, charts, and images can be used to identify patterns in data.</td></tr>
<tr><td>Clarifications and Content Limits</td><td colspan="3">Clarification Statement
• Emphasis is on finding patterns of changes in the level of complexity of anatomical structures in organisms and the chronological order of fossil appearance in the rock layers.

Content Limits
• Does not include: genetic analysis, comparisons of fossils to extant organisms, embryological evidence, genetic variation, inheritance, selective pressures.
• <u>Students do not need to know</u>: the names of individual species/genera or intervals of geological time, taxonomy, processes of fossil formation.</td></tr>
<tr><td>Science Vocabulary Students are Expected to Know</td><td colspan="3">Sedimentary rock, volcanic rock, radioactive dating, mineral, extinct, unicellular, multicellular, organelles, nucleus, ancestor, ancestry, species, evolve, anatomical.</td></tr>
<tr><td>Science Vocabulary Students are Not Expected to Know</td><td colspan="3">Cladogram, phylogenetics, phylogenetic systematics, phylum/phyla, class, order, family, genus/genera, homologous, analogous, divergent, convergent, prokaryote, eukaryote.</td></tr>
<tr><td colspan="4" align="center">Phenomena</td></tr>
<tr><td>Context/ Phenomena</td><td colspan="3">For this performance expectation the phenomena are sets of data. These are the observed facts that the kids will look at to discover patterns. Below, we enumerate some of the patterns that might comprise the data sets (phenomena) to be analyzed.

Stimuli might commonly include one or more geological column, data on what fossils are found in that (those) column(s), and the characteristics of those fossils. When more than one column is to be used in the analysis, sufficient data are given to anchor the ages of one or more key strata. Students would set out to identify and articulate patterns in the data.

Patterns that describe the data sets for MS-LS4-1:
• The first feather-like structures, associated with dinosaurs, appear in the fossil record close to 200 million years ago. Over the next 50 million years, a great variety of dinosaurs and true birds appeared, showing a great variety of feathers.</td></tr>
</table>

	• In North America, in the late C, a diverse assemblage of fossils is found. In the early Tertiary, there are far fewer types of fossils. • Prior to 542 million years ago, the fossil record shows relatively simple organisms without much variation. Layers in the fossil record between 542 million years ago to 476 million years ago shows the Cambrian Explosion—a time of significant evolution of animals, beginning with trilobites and ending with vertebrate fish. The Cambrian Explosion closed with a major extinction. • 525-year-old rock layers contain the earliest vertebrate fossils, which are of fish. These fossil fish had a cartilage skull with no jaw, and lacked a vertebral column. Fossils in 450-million-year-old rocks include vertebrate fish with a cartilage jaw and vertebral column. 400-year-old rocks include fish with skulls that include jaws and vertebrates made of bone.

This Performance Expectation and associated Evidence Statements support the following Task Demands.
Task Demands
1. Organize and/or arrange (e.g., using illustrations and/or labels) data that document patterns of change in the fossil record related to changes in anatomical structures or organism appearance/disappearance.
2. Generate/construct graphs, tables, or assemblages of illustrations and/or labels of data that document patterns of change in the fossil record related to changes in anatomical structures or organism appearance/disappearance. This may include sorting out distractors.
3. Determine or describe evidence that supports data on the timing of a mass extinction event, emergence/extinction of a new species/trait, and/or patterns of changes in biodiversity and organism complexity over time.
4. Identify/describe/illustrate/assemble sequences over time describing changes in characteristics of organisms, the diversity of the characteristics, the diversity of organisms, or the relative frequencies of the characteristics. This may include selecting a pattern from a list.

<table>
<tr><td>Performance Expectation</td><td colspan="3">MS-LS4-2
Apply scientific ideas to construct an explanation for the anatomical similarities and differences among modern organisms and between modern and fossil organisms to infer evolutionary relationships.</td></tr>
<tr><td>Dimensions</td><td>Constructing Explanations and Designing Solutions
• Apply scientific ideas to construct an explanation for real-world phenomena, examples, or events.</td><td>LS4.A: Evidence of Common Ancestry and Diversity
• Anatomical similarities and differences among organisms living today, and between contemporary organisms and those in the fossil record, enable the reconstruction of evolutionary history and the inference of lines of evolutionary descent.</td><td>Patterns
• Patterns can be used to identify cause and effect relationships.</td></tr>
<tr><td>Clarifications and Content Limits</td><td colspan="3">Clarification Statements
• Emphasis is on explanation of the relationships among organisms in terms of similarity or differences of the gross appearance of anatomical structures.
• Emphasis is on using anatomical similarities and differences to infer relationships among different modern organisms.
• Emphasis is on understanding that the changes over time in the anatomical features seen in fossil records can be used to infer relationships between extinct organisms to living organisms.
• Emphasis is on understanding that organisms that share a pattern of anatomical features are likely to be more closely related than are organisms that do not share a pattern of anatomical features.

Content Limits
• <u>Students do not need to know:</u> name of specific fossil species; knowledge of specific fossils or anatomical features; genetic variation, process of fossil formation, knowledge of geologic time periods; knowledge of rock layer; relationship between fossils and age of rock layers; molecular homology (similarities in DNA, RNA, and protein sequence).</td></tr>
<tr><td>Science Vocabulary Students are Expected to Know</td><td colspan="3">Homologous, analogous, diversity, extinction, radioactive dating, mineral, extinct, unicellular, multicellular, organelles, ancestry, species, evolve</td></tr>
<tr><td>Science Vocabulary Students are Not Expected to Know</td><td colspan="3">Cladogram, phylogenetic tree, dichotomous tree, phylum/phyla, class, order, family, genus/genera, divergent, convergent, prokaryote, eukaryote, types of rock (sedimentary, volcanic rock, igneous, metamorphic), embryology.</td></tr>
<tr><td colspan="4">Phenomena</td></tr>
<tr><td>Context/ Phenomena</td><td colspan="3">Some example phenomena for MS-LS4-2:
• Bats and frogs have forelimbs that look very different, but have similar bones and overall structure.
• Comparing the skull bones of the modern-day whale to the fossilized skulls of Dorudon and Pakicetus, shows a pattern in the position of the nostril as these organisms changed over millions of years.
• Wings are structures that allow most birds to fly, except penguins, which have wings but cannot fly.</td></tr>
</table>

	• Modern-day whales live in the ocean but have small hind-legs.
This Performance Expectation and associated Evidence Statements support the following Task Demands.	
Task Demands	
1. Articulate, describe, illustrate, or select the relationships, interactions, and/or processes to be explained. This may entail sorting relevant from irrelevant information or features.	
2. Express or complete a causal chain explaining how homologous structures show common ancestry and analogous structures show common function. This may include indicating directions of causality in an incomplete model, such as a flow chart or diagram, or completing cause and effect chains.*	
3. Identify evidence supporting the inference of causation that is expressed in a causal chain.	
4. Describe, identify, and/or select information needed to support an explanation.	

*denotes those task demands which are deemed appropriate for use in stand-alone item development

<table>
<tr><td>Performance Expectation</td><td colspan="3">MS-LS4-3
Analyze displays of pictorial data to compare patterns of similarities in embryological development across multiple species to identify relationships not evident in the fully formed anatomy.</td></tr>
<tr><td>Dimensions</td><td>Analyzing and Interpreting Data
• Analyze displays of data to identify linear and nonlinear relationships.</td><td>LS4.A: Evidence of Common Ancestry and Diversity
• Comparison of the embryological development of different species also reveals similarities that show relationships not evident in the fully formed anatomy.</td><td>Patterns
• Graphs, charts, and images can be used to identify patterns in data.</td></tr>
<tr><td>Clarifications and Content Limits</td><td colspan="3">Clarification Statements
• Emphasis is on inferring general patterns of relatedness among embryos of different organisms by comparing the macroscopic of diagrams or pictures.

Content Limits
• Assessment of comparisons is limited to observable (with the naked eye) appearances of anatomical structures in embryological development.</td></tr>
<tr><td>Science Vocabulary Students are Expected to Know</td><td colspan="3">Species, mammal, reproduce, mitosis, meiosis, body structure, limb, fetus, organ, tissues, cells.</td></tr>
<tr><td>Science Vocabulary Students are Not Expected to Know</td><td colspan="3">Placenta, homologous structures, external/internal fertilization, zygote, differentiation, gamete, blastula, mesoderm, endoderm, ectoderm, notochord.</td></tr>
<tr><td colspan="4">Phenomena</td></tr>
<tr><td>Context/ Phenomena</td><td colspan="3">For this performance expectation, the data will consist of pictures, diagrams, etc. Students will be challenged to find patterns and similarities.

Some example phenomena for MS-LS4-3:
• Early mammal embryos and early fish embryos both contain gill slits. In fish embryos, these gill slits develop into gills. In human embryos, the gill slits disappear before birth.
• The embryos of chickens, humans, and koalas have tails, and muscles to move the tails. However, as the embryos develop, the tails disappear.
• The limb buds of early bird embryos are very similar to the limb buds of early human embryos. The limb buds of the bird embryos become wings, while the limb buds of human embryos become arms.
• The early embryos of fish, birds, rabbits, and humans all have two-chambered hearts.</td></tr>
<tr><td colspan="4">This Performance Expectation and associated Evidence Statements support the following Task Demands.</td></tr>
<tr><td colspan="4">Task Demands</td></tr>
<tr><td colspan="4">1. Summarize data to highlight trends, patterns, or correlations in the similarities or differences of the embryonic development of different species.</td></tr>
</table>

2. Use relationships identified in the patterns of embryology data to predict the relatedness of different species.
3. Construct a statement that can potentially explain the observed trends or relationships in embryology data.
4. Identify patterns or evidence in the data that support inferences about the development of different species.
5. Identify additional information needed to support or challenge inferences based on identified patterns.

<table>
<tr><td>Performance Expectation</td><td colspan="3">MS-LS4-4
Construct an explanation based on evidence that describes how genetic variations of traits in a population increase some individuals’ probability of surviving and reproducing in a specific environment.</td></tr>
<tr><td>Dimensions</td><td>Constructing Explanations and Designing Solutions
• Construct an explanation that includes qualitative or quantitative relationships between variables that describe phenomena.</td><td>LS4.B: Natural Selection
• Natural selection leads to the predominance of certain traits in a population, and the suppression of others.</td><td>Cause and Effect
• Phenomena may have more than one cause, and some cause and effect relationships in systems can only be described using probability.</td></tr>
<tr><td>Clarifications and Content Limits</td><td colspan="3">Clarification Statements
• Emphasis is on using simple probability statements and proportional reasoning to construct explanations.
• Emphasize the use of proportional reasoning to support explanations of trends in changes to populations over time.
• Examples could include camouflage, variation of body shape, speed and agility, or drought tolerance.

Content Limits
• <u>Students do not need to know</u>: dominant/recessive traits, modes of inheritance (polygenic, sex-linked, etc.).</td></tr>
<tr><td>Science Vocabulary Students are Expected to Know</td><td colspan="3">Diversity, trend, predation, abundance, evolve/evolution, allele, sexual reproduction, beneficial, probability, distribution, adaptation, adaptive characteristics, frequency, DNA, dominant traits, recessive traits,</td></tr>
<tr><td>Science Vocabulary Students are Not Expected to Know</td><td colspan="3">Gene expression, polygenic traits, sex-linked traits, mutation, advantageous, cline, microevolution, gene pool, genetic drift, founder effect, bottleneck effect, gene flow, relative fitness.</td></tr>
<tr><td colspan="4">Phenomena</td></tr>
<tr><td>Context/ Phenomena</td><td colspan="3">Some example phenomena for MS-LS4-4:
• The orchid mantis attracts pollinators of the orchid as prey.
• In New Mexico, the rock pocket mice found in dark, rocky areas of the Valley of Fire all have dark fur.
• Male frigate birds with larger red pouches are more likely to find a mate.
• Some Staphylococus aureus bacteria are able to survive following treatment with the antibiotic methicillin.</td></tr>
<tr><td colspan="4">This Performance Expectation and associated Evidence Statements support the following Task Demands.</td></tr>
<tr><td colspan="4">Task Demands</td></tr>
<tr><td colspan="4">1. Describe or select the relationships, interactions, or processes to be explained. This may entail sorting relevant from irrelevant information or features.</td></tr>
<tr><td colspan="4">2. Complete a causal chain explaining how genetic variation affects the probability of survival and reproduction. This may include indicating directions of causality in a flow chart, diagram, or cause and effect chain.</td></tr>
</table>

3. Identify evidence supporting the role of genetic variation in determining the probability of survival and reproduction of an organism.
4. Predict changes in the frequency of a trait, given a change in the environment.
5. Identify the information needed to support an explanation for how genetic variation affects the rate of survival and reproduction.

<table>
<tr><td>Performance Expectation</td><td colspan="3">MS-LS4-5
Gather and synthesize information about technologies that have changed the way humans influence the inheritance of desired traits in organisms.</td></tr>
<tr><td>Dimensions</td><td>Obtaining, Evaluating, and Communicating
• Gather, read, and synthesize information from multiple appropriate sources and assess the credibility, accuracy, and possible bias of each publication and method used, and describe how they are supported or not supported by evidence.</td><td>LS4.B: Natural Selection
• In artificial selection, humans have the capacity to influence certain characteristics of organisms by selective breeding. One can choose desired parental traits determined by genes, which are then passed on to offspring.</td><td>Cause and Effect
• Phenomena may have more than one cause, and some cause and effect relationships in systems can only be described using probability.</td></tr>
<tr><td>Clarifications and Content Limits</td><td colspan="3">Clarification Statements
• Emphasis is on synthesizing information from reliable sources about the influence of humans on genetic outcomes in artificial selection (such as genetic modification, animal husbandry, and gene therapy) and on the impacts these technologies have on society as well as the technologies leading to these scientific discoveries.

Content Limits
• <u>Students do not need to know</u>: overlapping DNA sequences, Hardy-Weinberg calculations, biodiversity, mechanisms of gene transfer, dominant/recessive genes.</td></tr>
<tr><td>Science Vocabulary Students are Expected to Know</td><td colspan="3">Natural selection, artificial selection, evolution, adaptation, resources, reproduction, offspring, breeding, genetic engineering, DNA, cloning, inherit, hereditary, proteins.</td></tr>
<tr><td>Science Vocabulary Students are Not Expected to Know</td><td colspan="3">Chromosomes, genetic variation, genetic combination, meiosis, mitosis, replications, mutations, gene regulation, allele, RNA sequences, amino acid sequences.</td></tr>
<tr><td colspan="4" align="center">Phenomena</td></tr>
<tr><td>Context/ Phenomena</td><td colspan="3">Some example phenomena for MS-LS4-5:
• There is no wild plant that looks like modern corn (soft starchy kernels lined up in a row).
• Farmers isolated wild cabbage plants to create a variety of vegetables, including broccoli and kale. The wild cabbage plants were selected for their different flavors, textures, leaves, and flowers.
• Scientists are currently working to breed sheep that do not burp in order to reduce methane emission.
• Scientists want to breed strong and more resistant bees that won’t be damaged by disease and other parasites.</td></tr>
<tr><td colspan="4" align="center">This Performance Expectation and associated Evidence Statements support the following Task Demands.</td></tr>
<tr><td colspan="4" align="center">Task Demands</td></tr>
<tr><td colspan="4">1. Generate or construct tables or assemblages of data that document the similarities and differences between traditional and modern gene selection.</td></tr>
</table>

2. Organize and/or arrange data of the success rates of different methods to highlight trends or patterns in genetic modification.
3. Use relationships identified in the data to predict the best gene selection method to use in a given situation.
4. Identify, among distractors, the potential real-world uses of this data.

Performance Expectation	**MS-LS4-6** Use mathematical representations to support explanations of how natural selection may lead to increases and decreases of specific traits in populations over time.		
Dimensions	**Using Mathematics and Computational Thinking** • Use mathematical representations to support scientific conclusions and design solutions.	**LS4.C: Adaptation** • Adaptation by natural selection acting over generations is one important process by which species change over time in response to changes in environmental conditions. Traits that support successful survival and reproduction in the new environment become more common; those that do not become less common. Thus, the distribution of traits in a population changes.	**Cause and Effect** • Phenomena may have more than one cause, and some cause and effect relationships in systems can only be described using probability.
Clarifications and Content Limits	**Clarification Statement** • Emphasis is on using mathematical models, probability statements, and proportional reasoning to support explanations of trends in changes to populations over time. **Content Limits** • Math can include measures of central tendency, basic operations that can be calculated without a calculator, and basic graphical analysis (bar chart, pie chart, scatter plot, box and whisker plot, line chart). • Students aren't expected to know the mechanisms of genetic inheritance or mutation. • Assessment does not include Hardy-Weinberg calculations. • Assessment does not include other mechanisms of evolution (genetic drift, co-evolution, gene flow, etc.) • <u>Students do not need to know</u>: Alleles, DNA sequences, anatomical structures, embryonic development, gene frequency, morphology, speciation.		
Science Vocabulary Students are Expected to Know	Climate, evolution, inherit, generation, species, genus, reproduction, distribution, ratio .		
Science Vocabulary Students are Not Expected to Know	Morphology, genetic variance, proliferation, biotic/abiotic.		
Phenomena			
Context/ Phenomena	Some example phenomena for MS-LS4-6: • Some bacteria are killed by a certain antibiotic while other bacteria are immune to it. After the antibiotic is used once, bacteria die. The next time the antibiotic is used, there are many bacteria left. • The Sandhills in Nebraska used to be covered in dark-colored soil. Most deer mice living in this area had dark-colored fur coats, while others had light-colored fur coats. Over time, the Sandhills were covered in light-colored sand. After many years, the population of deer mice had mostly light-colored fur coats. This will be presented as data.		

	• In the Galapagos Islands, there are finches with thin, small beaks that eat small, soft seeds. There also finches with thick, large beaks that eat larger hard and dry seeds. A drought period in 1977 affected the plant life on the islands, greatly reducing the number of small, soft seeds. The next year, there were far more large-beaked birds than small-beaked birds.

This Performance Expectation and associated Evidence Statements support the following Task Demands.
Task Demands
1. Make simple calculations using given data to calculate or estimate changes in the prevalence of specific traits over time.
2. Illustrate, graph, or calculate the prevalence of specific traits passed on in observed populations under varying conditions, from given data. The data may be ordinal and the calculations may be representations of trends or propensities.
3. Calculate or estimate properties or relationships of the changes in the distribution of traits among a population under varying conditions, based on data from one or more sources.
4. Compile, from given information, the data needed for a particular inference about the relationship between changes in the environment and changes in the traits of a population.
5. Use mathematical representations and/or computational representations (such as trends, averages, histograms, graphs, spreadsheets) to identify relationships in the data.
6. Use mathematical representations and/or computational representations (such as trends, averages, histograms, graphs, spreadsheets) to explain the influence that natural selection has had on the presence of specific traits in a population over time.

<table>
<tr><td>Performance Expectation</td><td colspan="3">MS-ESS1-1
Develop and use a model of the Earth-sun-moon system to describe the cyclic patterns of lunar phases, eclipses of the sun and moon, and the seasons.</td></tr>
<tr><td>Dimensions</td><td>Developing and Using Models
• Develop and use a model to describe phenomena.</td><td>ESS1.A The Universe and Its Stars
• Patterns of the apparent motion of the sun, the moon, and stars in the sky can be observed, described, predicted, and explained with models.

ESS1.B Earth and the Solar System
• This model of the solar system can explain eclipses of the sun and the moon. Earth’s spin axis is fixed in direction over the short term but tilted relative to its orbit around the sun. The seasons are a result of that tilt and are caused by the differential intensity of sunlight on different areas of Earth across the year.</td><td>Patterns
• Patterns can be used to identify cause and effect relationships.</td></tr>
<tr><td>Clarifications and Content Limits</td><td colspan="3">Clarification Statements
• Examples of models can be physical, graphical, or conceptual.

Content Limits
• <u>Students do not need to know</u> Earth’s exact tilt; sidereal and synodic periods; umbra and penumbra (the term “shadow” should be used); times of moonrise and moonset; precession; exact dates of equinoxes and solstices (but knowledge of the months in which they occur is reasonable to assess).</td></tr>
<tr><td>Science Vocabulary Students are Expected to Know</td><td colspan="3">Shadow, orbit, axis, planet, satellite, full moon, new moon, half moon</td></tr>
<tr><td>Science Vocabulary Students are Not Expected to Know</td><td colspan="3">Perigee, apogee, sidereal period, sidereal month, synodic period, synodic month, umbra, penumbra, precession, equinox, solstice, ecliptic, waxing, waning, gibbous, first quarter moon, last quarter moon</td></tr>
<tr><td colspan="4">Phenomena</td></tr>
<tr><td>Context/ Phenomena</td><td colspan="3">Some example phenomena for MS-ESS1-1:
• When observed from Earth over the course of a month, the appearance of the moon changes.
• A full moon occurs in every calendar month. However, an eclipse of the moon does not occur in every calendar month.
• A new moon occurs in every calendar month. However, a total eclipse of the sun is a rare event.
• In the northern hemisphere, July is a summer month. In the southern hemisphere, July is a winter month.</td></tr>
<tr><td colspan="4">This Performance Expectation and associated Evidence Statements support the following Task Demands.</td></tr>
<tr><td colspan="4">Task Demands</td></tr>
</table>

7. Select or identify from a collection of potential model components, including distractors, components needed for a model that can explain lunar phases, eclipses of the sun, eclipses of the moon, *or* seasons on Earth. Components might include the sun, moon, Earth, solar energy, the moon’s orbital trace, Earth’s orbital trace, the angle of the moon’s orbital trace, the angle of Earth’s orbital trace, Earth’s axis, or the tilt of Earth’s axis.
8. Assemble or complete, from a collection of potential model components, an illustration or flow chart that is capable of representing the causes of lunar phases, eclipses of the sun, eclipses of the moon, *or* seasons on Earth. This <u>does not</u> include labeling a simple diagram of the Earth-sun-moon system.
9. Describe, select, or identify the relationships among components of a model that can explain lunar phases, eclipses of the sun, eclipses of the moon, *or* seasons on Earth. Components might include the sun, moon, Earth, solar energy, the moon’s orbital trace, Earth’s orbital trace, the angle of the moon’s orbital trace, the angle of Earth’s orbital trace, Earth’s axis, or the tilt of Earth’s axis.
10. Manipulate the components of a model to demonstrate how the relationships among the sun, the moon, Earth, and solar energy change to result in lunar phases, eclipses of the sun, eclipses of the moon, *or* seasons on Earth. *****
11. Make predictions about the effects of changes in the relationships among the sun, the moon, Earth, and solar energy as they relate to lunar phases, eclipses of the sun, eclipses of the moon, *or* seasons on Earth. Predictions can be made by manipulating model components, completing illustrations, or selecting from lists with distractors. *****
12. Identify missing components, relationships, or other limitations of a model that can explain lunar phases, eclipses of the sun, eclipses of the moon, *or* seasons on Earth.

*denotes those task demands which are deemed appropriate for use in stand-alone item development

<table>
<tr><td>Performance Expectation</td><td colspan="3">MS-ESS1-2
Develop and use a model to describe the role of gravity in the motions within galaxies and the solar system.</td></tr>
<tr><td>Dimensions</td><td>Developing and Using Models
• Develop and use a model to describe phenomena.</td><td>ESS1.A: The Universe and Its Stars
• Earth and its solar system are part of the Milky Way galaxy, which is one of many galaxies in the universe.

ESS1.B: Earth and the Solar System
• The solar system consists of the sun and a collection of objects, including planets, their moons, and asteroids that are held in orbit around the sun by its gravitational pull on them.
• The solar system appears to have formed from a disk of dust and gas, drawn together by gravity.</td><td>Systems and System Models
• Models can be used to represent systems and the interactions in a system.</td></tr>
<tr><td>Clarifications and Content Limits</td><td colspan="3">Clarification Statements
• Emphasis for the model is on gravity as the force that holds together the solar system and Milky Way galaxy, and controls orbital motions within them.
• Examples of models can be physical (such as the analogy of distance along a football field or computer visualizations of elliptical orbits) or conceptual (such as mathematical proportions relative to the size of familiar objects such as students' school or state).
• Focus should be on qualitative comparisons, not quantitative.

Content Limits
• Assessment does not include Kepler's Laws of orbital motion or the apparent retrograde motion of the planets as viewed from Earth.
• Assessment does not include specific facts about any planets or moons.</td></tr>
<tr><td>Science Vocabulary Students are Expected to Know</td><td colspan="3">Inertia, force, mass, weight, orbit, names of planets</td></tr>
<tr><td>Science Vocabulary Students are Not Expected to Know</td><td colspan="3">Names of specific moons, names of space shuttles, moment of inertia, Kepler's laws of planetary motion, black hole</td></tr>
<tr><td colspan="4">Phenomena</td></tr>
<tr><td>Context/ Phenomena</td><td colspan="3">Some example phenomena for MS-ESS1-2:
• Satellites orbit Earth but can fall out of orbit (Skylab, UART satellite).
• Halley's Comet can be seen as it travels past Earth every 75–76 years.
• Rings are present around some planets.
• Mars has two moons, Phobos and Deimos, which orbit the planet.</td></tr>
<tr><td colspan="4">This Performance Expectation and associated Evidence Statements support the following Task Demands.</td></tr>
<tr><td colspan="4">Task Demands</td></tr>
</table>

1. Select or identify from a collection of potential model components, including distractors, the components needed for a model that describes the role of gravity in celestial bodies.
2. Assemble or complete, from a collection of potential model components, an illustration, diagram or description that is capable of representing forces and their influences on the motion of celestial bodies and/or man-made objects in orbit. This does not include labeling an existing diagram.
3. Describe, select or identify the relationships among components of a model that can explain the role of gravity in the motions of galaxies and the solar system. Components might include the sun, the moon, Earth, Milky Way galaxy, other planets and their moons.
4. Manipulate the components of a model to demonstrate how the relationships among the sun, the Earth, the moon, planets in the solar system, and galaxies change the resulting gravitational force between/or motions of those bodies.*
5. Make predictions about the effects of changes in mass/distance/how fast an object travels in a given model on other objects in the system. Predictions can be based on manipulating model components, completing illustrations, or selecting from a list including distractors.
6. Identify missing components, relationships, or other limitations of a model that can explain the role of gravity.

*denotes those task demands which are deemed appropriate for use in stand-alone item development

<table>
<tr><td>Performance Expectation</td><td colspan="3">MS ESS1-3
Analyze and interpret data to determine scale properties of objects in the solar system.</td></tr>
<tr><td>Dimensions</td><td>Analyzing and Interpreting Data
• Analyze and interpret data to determine similarities and differences in findings.</td><td>ESS1.B: Earth and the Solar System
• The solar system consists of the sun and a collection of objects, including planets, their moons, and asteroids that are held in orbit around the sun by its gravitational pull on them.</td><td>Scale, Proportion, and Quantity
• Time, space, and energy phenomena can be observed at various scales using models to study systems that are too large or too small.</td></tr>
<tr><td>Clarifications and Content Limits</td><td colspan="3">Clarification Statements
• Emphasis is on the analysis of data from Earth-based instruments, space-based telescopes and spacecraft to determine similarities and differences among solar system objects.
• Examples of scale properties include the sizes of an object’s layers (such as crust and atmosphere), surface features (such as volcanoes), and orbital radius.
• Examples of data include statistical information, drawings and photographs, and models.

Content Limits
• Assessment does not include recalling facts about properties of the planets and other solar system bodies.
• <u>Students do not need to know</u>: Facts about properties of the planets and other solar system bodies, scientific notation.</td></tr>
<tr><td>Science Vocabulary Students are Expected to Know</td><td colspan="3">Satellite, terrestrial planet, gas giant, planetary rings, dwarf planet, sun, inner planet, outer planet, comet,</td></tr>
<tr><td>Science Vocabulary Students are Not Expected to Know</td><td colspan="3">Density, ecliptic, solar wind, interstellar medium, main sequence, synchronous rotation, protostar, protoplanetary disc, accretion.</td></tr>
<tr><td colspan="4" align="center">Phenomena</td></tr>
<tr><td>Context/ Phenomena</td><td colspan="3">The phenomena for this performance expectation are the given data. Samples of phenomena should describe the data set(s) to be given in terms of patterns or relationships to be found in the data, and the columns and rows of a hypothetical table presenting the data, even if the presentation is not tabular. The description of the phenomenon should describe the presentation format of the data (e.g., maps, tables, graphs, etc).

Some example phenomena for MS-ESS1-3:
• Four of Jupiter’s moons can be clearly seen through a small telescope under low magnification. These moons appear as tiny dots arranged around Jupiter.
• Close-up pictures from the New Horizons mission provided new evidence about the dwarf planet, Pluto, which was not able to be gathered by distant observations and calculations (surface features, scale).
• The sun and the moon appear as approximately the same size in the sky, but the sun is vastly larger than the moon (scale).
• Even though the moon is infinitesimally smaller than the sun, the entire sun is blocked from view on Earth during a solar eclipse (scale).</td></tr>
</table>

This Performance Expectation and associated Evidence Statements support the following Task Demands.	
Task Demands	
1. Make simple calculations using given data to estimate the properties (e.g., mass, surface temp., diameter) and locations of different solar system objects relative to a given reference point/object.	
2. Illustrate, graph, or identify relevant features or data that can be used to estimate properties of objects or relationships in our solar system.	
3. Calculate, estimate or identify properties of objects or relationships among objects in the solar system, based on data from one or more sources.*	
4. Compile, from given information, the data needed for a particular inference about scale or other properties of an object.	
5. Given a partial model of objects in the solar system, identify objects or relationships that can be represented in the model or the reasons why they cannot be represented in the model.	

*denotes those task demands which are deemed appropriate for use in stand-alone item development

Performance Expectation	**MS-ESS1-4** Construct a scientific explanation based on evidence from rock strata for how the geologic timescale is used to organize Earth's 4.6-billion-year-old history.		
Dimensions	**Constructing Explanations and Designing Solutions** • Construct a scientific explanation based on valid and reliable evidence obtained from sources (including students' own experiments) and the assumption that theories and laws that describe the natural world operate today as they did in the past and will continue to do so in the future.	**ESS1.C: The History of Planet Earth** • The geological time scale interpreted from rock strata provides a way to organize Earth's history. Analyses of rock strata and the fossil record provide only relative dates, not an absolute scale.	**Scale, Proportion and Quantity** • Time, space, and energy phenomena can be observed at various scales, using models to study systems that are too large or too small.
Clarifications and Content Limits	**Clarification Statements** • Emphasis is on how analyses of rock formations and the fossils they contain are used to establish relative ages of major events in Earth's history. • Example of Earth's major events could range from being geologically recent (e.g., the most recent glacial period or the earliest fossils of Homo sapiens) to geologically very old (e.g., the formation of Earth or the earliest evidence of life). • Examples can include the formation of mountain chains and ocean basins, the evolution or extinction of particular living organisms, or significant instances of volcanic eruptions. **Content Limits** • Assessment does not include recalling the names of specific periods and epochs or events within them.		
Science Vocabulary Students are Expected to Know	Erosion, weathering, ancient, prehistoric, layer, formation, mineral, sedimentary, sediment, metamorphic, volcanic, superposition, cross-cutting, fault, fold, geology, geological		
Science Vocabulary Students are Not Expected to Know	Radioactive dating, bio-geology, geobiology, relative dating, numerical dating, absolute dating, carbon dating, radiometric dating, igneous, stratigraphy, biostratigraphy, chronostratigraphy, sequence stratigraphy, bed, lamina, paleoenvironment, paleoecology, paleomagnetic		
Phenomena			
Context/ Phenomena	Some example phenomena for MS-ESS1-4: • A very distinct clay layer tops the Hell Creek Formation in Montana. Below this layer, the Hell Creek is rich in dinosaur fossils; above the layer, no dinosaurs are found. • The landscape of Cape Cod, Massachusetts, is almost entirely small hills of sand and gravel. However, a hole drilled 500 feet into the ground will hit hard metamorphic rock. • In Box Canyon in Ouray, Colorado, metamorphic rocks that are standing vertical are capped by sedimentary rocks that are lying flat. • The St. Peter Sandstone is a very white sandstone rock layer exposed in many places in the midwestern United States. The St. Peter is very uniform in appearance but the rock layer sits on top of different kinds of rocks in the North than it does in Missouri.		
This Performance Expectation and associated Evidence Statements support the following Task Demands.			

Task Demands
1. Organize and/or arrange (e.g., using illustrations and/or labels, including taken from or added to, stratigraphic columns and/or geologic maps), or summarize, data/information so as to highlight trends, patterns, or correlations in paleoenvironmental changes, geological events/processes, and/or the appearance or disappearance in the record of specific organisms.*****
2. Generate/construct graphs, tables, or assemblages of illustrations, and/or labels of data/information that document patterns, trends, or correlations in how rock types and included fossils change over geologic time, recording different events and paleo environments. This may include sorting out distractors.*****
3. Use relationships identified in the data/information to hypothesize the relative age of specific rock layers, formations, or fossils, in a stratigraphic column or on a geologic map.*****
4. Identify patterns or evidence in the data/information that support inferences about what the paleoenvironment was like during time intervals represented in a stratigraphic column or on a geologic map.
5. Describe, identify, and/or select information needed to support an explanation.

*denotes those task demands which are deemed appropriate for use in stand-alone item development. 2/3 of these TDs should be combined and used when developing a stand-alone item.

<table>
<tr><td>Performance Expectation</td><td colspan="3">MS-ESS2-1
Develop a model to describe the cycling of Earth’s materials and the flow of energy that drives this process.</td></tr>
<tr><td>Dimensions</td><td>Developing and Using Models
• Develop and use a model to describe the phenomena.</td><td>ESS2.A: Earth’s Materials and Systems
• All Earth processes are the result of energy flowing and matter cycling within and among the planet’s systems. This energy is derived from the sun and Earth’s hot interior. The energy that flows and matter that cycles produce chemical and physical changes in Earth’s materials and living organisms.</td><td>Stability and Change
• Explanations of stability and change in natural or designed systems can be contructed by examining the changes over time and processes at different scales, including the atomic scale.</td></tr>
<tr><td>Clarifications and Content Limits</td><td colspan="3">Clarification Statements
• Emphasis is on the processes of melting, crystallization, weathering, sedimentation, and deformation, which act together to form minerals and rocks through the cycling of Earth’s matter.

Content Limits
• Assessment does not include the identification and naming of minerals.
• <u>Students do not need to know</u>: specific processes of chemical or biogeochemical weathering; rock phase diagrams; mineral stability diagrams; mineral weathering orders; mineral crystallization orders (e.g., Bowen’s Reaction Series); mineral metamorphism orders/temperatures/pressures/stabilities; rock metamorphism zones; specific processes that drive the tectonic engine (e.g., slab pull; ridge push).</td></tr>
<tr><td>Science Vocabulary Students are Expected to Know</td><td colspan="3">Collide, heat conduction, transform, transport, heat transfer, heat radiation, thermal energy, heat convection, precipitation, volcanic eruption, chemical, weathering, erosion, sediment, deposition, rock cycle, ice wedge, fault, fold, igneous rock, metamorphic rock, sedimentary rock, volcanic rock, plate tectonics, crust, mantle, outer core, inner core.</td></tr>
<tr><td>Science Vocabulary Students are Not Expected to Know</td><td colspan="3">Biogeology, geobiology, geochemistry, biogeochemistry, rock sequence, convection current, mountain building, geochemical cycle, tectonic uplift, accretionary wedge, accretionary prism.</td></tr>
<tr><td colspan="4">Phenomena</td></tr>
<tr><td>Context/ Phenomena</td><td colspan="3">Some example phenomena for MS-ESS2-1:
• Lava from an erupting volcano in Hawaii flows across a road. The molten material is so hot that it emits light. Several months later, the material covering the road is a hard, black rock.
• A mountain is capped by metamorphic rock. Many cracks crisscross the rock. Rainwater often fills the fractures, freezing when temperatures drop. Over the years, the fractures become wider.
• An exposure of bedded sandstone has been cut by a plug of igneous rock. Near the edges of the igneous rock, the sandstone is discolored and displays a different texture from the rest of the exposure.
• An exposure of sedimentary rock contains pieces of a metamorphic rock that is exposed several miles away.</td></tr>
<tr><td colspan="4">This Performance Expectation and associated Evidence Statements support the following Task Demands.</td></tr>
<tr><td colspan="4">Task Demands</td></tr>
</table>

1. Select or identify from a collection of potential model components, including distractors, the components needed to model the phenomenon. Components might include different rock types, processes that change one rock type into another, surface environments on Earth where these processes occur and where different rock types exist, and layers within Earth where these processes occur. Sources of energy (radiation, convection) that drive the cycling (but *not* the creation of) matter should also be included as components.
2. Assemble or complete, from a collection of potential model components, an illustration, virtual representation of a physical model, or flow chart that is capable of representing how energy (radiation, convection) drives processes that cycle (but do *not* create) matter on Earth. This *does not* include labeling an existing diagram.
3. Manipulate the components of a model to demonstrate the changes, properties, processes, and/or events that act to resultin the cycling of Earth's materials.
4. Make predictions about the effects of changes in the rock cycle. Predictions can be made by manipulating model components, completing illustrations, or selecting from lists with distractors.
5. Given models or diagrams of the rock cycle, identify different rock types and how they change in each scenario OR identify the properties of energy that cause Earth materials to cycle between different rock types.
6. Identify missing components, relationships, or other limitations of the model that can explain the cycling of Earth's materials.
7. Identify or select the relationships among components of a model that describe the relationship between energy and the cycling of matter that forms different types of rock, or explain the relationship between energy and the cycling of matter that forms different types of rock.

<table>
<tr><td>Performance Expectation</td><td colspan="3">MS-ESS2-2
Construct an explanation based on evidence for how geoscience processes have changed Earth's surface at varying time and spatial scales.</td></tr>
<tr><td>Dimensions</td><td>Constructing Explanations and Designing Solutions
• Construct a scientific explanation based on valid and reliable evidence obtained from sources (including the students' own experiments) and the assumption that theories and laws that describe nature operate today as they did in the past and will continue to do so in the future.</td><td>ESS2.A: Earth's Materials and Systems
• The planet's systems interact over scales that rage from microscopic to global in size, and they operate over fractions of a second to billions of years. These interactions have shaped Earth's history and will determine its future.

ESS2.C: The Roles of Water in Earth's Surface Processes
• Water's movements—both on the land and underground—cause weathering and erosion, which change the land's surface features and create underground formations.</td><td>Scale, Proportion, and Quantity
• Time, space, and energy phenomena can be observed at various scales using models to study systems that are too large or too small.</td></tr>
<tr><td>Clarifications and Content Limits</td><td colspan="3">Clarification Statements
• Emphasis is on how processes change Earth's surface at time and spatial scales that can be large (such as slow plate motions or the uplift of large mountain ranges) or small (such as rapid landslides or microscopic geochemical reactions), and how many geoscience processes (such as earthquakes, volcanoes, and meteor impacts) usually behave gradually but are punctuated by catastrophic events.
• Examples of geoscience processes include surface weathering and deposition by the movements of water, ice, and wind.
• Emphasis is on geoscience processes that shape local geographic features, where appropriate.

Content Limits
• Students are expected to know all of the components/processes of the rock cycle but not specific rock or mineral names.
• <u>Students do not need to know</u> Endogenic or exogenic systems, specific intervals of the Geological Time Scale by name, specific volcano types (shield, effusive, composite, etc.)/</td></tr>
<tr><td>Science Vocabulary Students are Expected to Know</td><td colspan="3">Earthquake, volcanic eruptions, core, crust, mantle, pressure, continent, erosion, weathering, magma, lava, igneous, sedimentary, metamorphic, mineral, meteor, crater, plate tectonics, continental drift, subduction zone, divergent boundary, convergent boundary, hot spot, fault, tsunami, hurricane, tornado, fracture, folding, compressing, sea floor spreading, layer, ridge, rock cycle, trench, plateau, slope, landslides, floods, caves</td></tr>
<tr><td>Science Vocabulary Students are Not Expected to Know</td><td colspan="3">Endogenic system, exogenic system, radiometric dating, original horizontality, superposition, uniformitarianism, primordial, epoch, eon, period, liquification, Mohorovicic discontinuity (Moho), seismic waves, seismograph, Richter scale, fumaroles, mofettes, solfataras, Caledonian era, Variscan era, Alpine era, massif, graben, monolith, monadnock, nappe system, isostasy, pluton, batholith, stratigraphy, lithification, evaporite, hydrothermal, relief, topography, continental shield, terrain, anticline, syncline, strike-slip fault, horst, orogenesis, tephra, caldera</td></tr>
<tr><td colspan="4">Phenomena</td></tr>
<tr><td>Context/ Phenomena</td><td colspan="3">Some example phenomena for MS-ESS2-2:</td></tr>
</table>

	• A hillside in Oregon experiences an intense rain storm. At the end of the storm, part of the hillside collapses, covering a road with mud and debris. • In Northern Arizona, there is a large circular depression. • In southeastern Pennsylvania, the landscape is dotted with a number of irregular holes that lead to caves. • When viewed from orbit, the Eastern coastline of South America and the Western Coast of Africa look as though they were joined together, similar to a jigsaw puzzle.
This Performance Expectation and associated Evidence Statements support the following Task Demands.	
Task Demands	
1. Articulate, describe, illustrate, or select the relationships, interactions, and/or processes to be explained. This may entail sorting relevant from irrelevant information or features.	
2. Express or complete a causal chain explaining how a given process(es) acts to modify Earth’s surface in the long term and/or short term. This may include indicating directions of causality in an incomplete model such as a flow chart or diagram, or completing cause and effect chains.	
3. Identify evidence supporting the inference of causation that is expressed in a causal chain for a process(es) that acts to modify Earth’s surface in the long term and/or short term.	
4. Use an explanation to predict the effect of the process on Earth’s surface, given a change in conditions (e.g., atmospheric, tectonic, geological, hydrologic).	
5. Describe, identify, and/or select information needed to support an explanation for how processes affect Earth’s surface over the short and/or long term.	

<table>
<tr><td>Performance Expectation</td><td colspan="3">MS-ESS2-3
Analyze and interpret data on the distribution of fossils and rocks, continental shapes, and seafloor structures to provide evidence of past plate motions.</td></tr>
<tr><td>Dimensions</td><td>Analyzing and Interpreting Data
• Analyze and interpret data to provide evidence for phenomena.</td><td>ESS1.C: The History of Planet Earth
• Tectonic processes continually generate new ocean sea floor at ridges and destroy old sea floor at trenches. (secondary)

ESS2.B: Plate Tectonics and Large-Scale System Interactions
• Maps of ancient land and water patterns, based on investigations of rocks and fossils, make clear how Earth’s plates have moved great distances, collided, and spread apart.</td><td>Patterns
• Patterns in rates of change and other numerical relationships can provide information about natural systems.</td></tr>
<tr><td>Clarifications and Content Limits</td><td colspan="3">Clarification Statements
• Examples of data include similarities of rock and fossil types on different continents, the shapes of continents (including continental shelves), and the locations of ocean structures (such as ridges fracture zones, and trenches).

Content Limits
• Paleomagnetic anomalies in oceanic and continental crust are not assessed.
• <u>Students do not need to know</u>: Specific chemical makeup of the crust, mantle, and core; specific rocks within major categories (e.g., basalt, amphibolite, granite); mineral crystallization orders (e.g., Bowen’s Reaction Series), mineral melt orders.</td></tr>
<tr><td>Science Vocabulary Students are Expected to Know</td><td colspan="3">Crust, mantle, core, convection, density, plate tectonics, earthquake, geosphere, element, continental, upwelling, convection, trench, subduction, ridge, volcanic, sedimentary, fault, extension, volcanic rock, sedimentary rock, metamorphic rock, ridge, hotspot.</td></tr>
<tr><td>Science Vocabulary Students are Not Expected to Know</td><td colspan="3">Block (as in fault), accretionary wedge, accretionary prism, mantle composition, stress (tectonic), strain (tectonic), normal fault, transform fault, thrust fault, reverse fault, foot wall, hanging wall, felsic, mafic, ultramafic.</td></tr>
<tr><td colspan="4">Phenomena</td></tr>
<tr><td>Context/ Phenomena</td><td colspan="3">Some example phenomena for MS-ESS2-3:
• There are volcanoes on all of the Hawaiian islands. But only volcanoes on the southeastern most island, Hawaii, are active today.
• Earthquakes are very commonly felt on the islands of Japan.
• The Atlantic coasts of South America and Africa appear to fit together like two jigsaw puzzle pieces. Identical fossils of certain plants and animals are preserved in rocks found along both coasts.
• Earthquakes are very rare in the State of Florida.</td></tr>
<tr><td colspan="4">This Performance Expectation and associated Evidence Statements support the following Task Demands.</td></tr>
<tr><td colspan="4">Task Demands</td></tr>
<tr><td colspan="4">1. Identify patterns or evidence in the data that supports conclusions about how the Earth’s plates have moved and interacted with each other (e.g., converged or diverged).*</td></tr>
</table>

2. Use relationships identified in the data to predict the locations of fossils, earthquakes, or volcanoes.
3. Illustrate, graph or identify relevant features or data that can be used to identify past plate motions or estimate the rate of change in tectonic processes.
4. Organize and/or arrange (e.g., using illustrations and/or labels), or summarize data to highlight trends, patterns, or correlations.*
5. Compile from given information, the data needed to identify a pattern in the rate of change or evidence of past plate motions.

*denotes those task demands which are deemed appropriate for use in stand-alone item development

<table>
<tr><td>Performance Expectation</td><td colspan="3">MS-ESS2-4
Develop a model to describe how the cycling of water through Earth’s systems is driven by energy from the sun, gravitational forces, and density.</td></tr>
<tr><td>Dimensions</td><td>Developing and Using Models
• Develop a model to describe unobservable mechanisms.</td><td>ESS2.C: The Roles of Water in Earth’s Surface Processes
• Water continually cycles among land, ocean, and atmosphere via transpiration, evaporation, condensation and crystallization, and precipitation, as well as downhill flows on land.
• Global movements of water and its changes in form are propelled by sunlight and gravity.</td><td>Energy and Matter
• Within a natural or designed system, the transfer of energy drives the motion and/or cycling of matter.</td></tr>
<tr><td>Clarifications and Content Limits</td><td colspan="3">Clarification Statements
• Examples of models can be conceptual or physical.
• Content emphasis is on the ways water changes its state as it moves through the multiple pathways of the hydrologic cycle.
• Practice emphasis is on developing a model and being able to explain reasoning behind choices made relative to the developing or changing of a model. While a few interactions can be about using the model, the focus should not be on using the model or designing an experiment using the model. Any stand-alone items written to this PE should be centered on the development of models.

Content Limits
• A quantitative understanding of the latent heats of vaporization and fusion is not assessed.
• <u>Students do not need to know</u>:
o Cloud types
o Types of aquifers and components of aquifers
o Concepts of subsurface water flow and transmissivity (e.g., permeability/porosity of the substrate and interactions with fluids; behaviors of subsurface fluids under confinement (both quantitatively and qualitatively).</td></tr>
<tr><td>Science Vocabulary Students are Expected to Know</td><td colspan="3">Precipitation, transpiration, evaporation, condensation, crystallization, density, runoff, temperature, air pressure, particle, atmosphere</td></tr>
<tr><td>Science Vocabulary Students are Not Expected to Know</td><td colspan="3">Hyporheic zone, aquifer, aquitard, aquiclude, subsurface flow, sublimation, vadose zone, unsaturated zone, water table, phreatic surface, capillary fringe, saturated zone, phreatic zone, drainage basin, watershed, porosity, permeability, transmissivity, recharge, recharge area, discharge, discharge area, potentiometric surface, hydraulic head, lithosphere, biosphere, hydrosphere, cryosphere</td></tr>
<tr><td colspan="4" align="center">Phenomena</td></tr>
<tr><td>Context/ Phenomena</td><td colspan="3">Some example phenomena for MS-ESS2-4:
• When driving over a bridge on a cool morning, you see fog over the river but not over the land.
• Morning fog and mist soon disappears after the sun rises on a clear day.
• The Blue Mountains have snow that melts (eventually) into the Columbia River to the John Day Dam
• In the Iowa cornfields in the summer, a dense dome of humidity forms over the cornfields.</td></tr>
</table>

This Performance Expectation and associated Evidence Statements support the following Task Demands.	
Task Demands	
1. Select or identify from a collection of potential model components including distractors, the components needed to model the model of evaporation, condensation, transpiration, precipitation or other behaviors of water molecules during the water cycle.	
2. Assemble or complete, from a collection of potential model components, an illustration or flow chart that is capable of representing the phenomenon. This does not include labeling an existing diagram.*	
3. Manipulate the components of a model to demonstrate the effects those adjustments would have on the behavior of water in the water molecules in the water cycle.*	
4. Make predictions about the effects of changes to the parts of the model. Predictions can be based on manipulating model components, completing illustrations, or selecting from a list with distractors.	
5. Identify missing components, relationships, or other limitations of the model.	
6. Describe, select, or identify the relationships among components of a model that describe or explains the phenomenon.	
7. Identify, describe or explain reasons for choosing components of a model of the water cycle.	

*denotes those task demands which are deemed appropriate for use in stand-alone item development

<table>
<tr><td>Performance Expectation</td><td colspan="3">MS-ESS2-5
Collect data to provide evidence for how the motions and complex interactions of air masses result in changes in weather conditions.</td></tr>
<tr><td>Dimensions</td><td>Planning and Carrying Out Investigations
• Collect data to produce data to serve as the basis for evidence to answer scientific questions or test design solutions under a range of conditions.</td><td>ESS2.C: The Roles of Water in Earth’s Surface Processes
• The complex patterns of the changes and the movement of water in the atmosphere, determined by winds, landforms, and ocean temperatures and currents, are major determinants of local weather patterns.

ESS2.D: Weather and Climate
• Because these patterns are so complex, weather can only be predicted probabilistically.</td><td>Cause and Effect
• Cause and effect relationships may be used to predict phenomena in natural or designed system.</td></tr>
<tr><td>Clarifications and Content Limits</td><td colspan="3">Clarification Statements
• Emphasis is on how air masses flow from regions of high pressure to low pressure, causing weather (defined by temperature, pressure, humidity, precipitation, and wind) at a fixed location to change over time, and how sudden changes in weather can result when different air masses collide.
• Emphasis is on how weather can be predicted within probabilistic ranges.
• Examples of data can be provided to students (such as weather maps, diagrams, and visualizations) or obtained through laboratory experiments (such as condensation).

Content Limits
• Assessment does not include recalling the names of cloud types or weather symbols used on weather maps or the reported diagrams from weather stations.
• Weather incidents internal to air masses are excluded because the focus is on the interfaces between large scale air masses.
• <u>Students do not need to know</u>: Names of the various types of clouds, weather symbols used on weather maps, weather symbols used on reports from weather stations. A legend will be included on weather maps.</td></tr>
<tr><td>Science Vocabulary Students are Expected to Know</td><td colspan="3">Density, temperature, pressure, humidity, precipitation, wind speed, wind direction, air mass, cold/warm front, condensation, evaporation, latitude, altitude, flow, thermometer, barometer, anemometer, dew point, stationary front, occluded front, warm front, cold front</td></tr>
<tr><td>Science Vocabulary Students are Not Expected to Know</td><td colspan="3">Horse latitudes, Tropic of Capricorn, Tropic of Cancer, cyclone, anticyclone, isobar, isotherm, pressure gradient, Coriolis force*, hygrometer and psychrometer (humidity meters), jet stream,
*Coriolis force IS covered in PE MS-ESS2-6, however.</td></tr>
<tr><td colspan="4">Phenomena</td></tr>
<tr><td>Context/ Phenomena</td><td colspan="3">Some example phenomena for MS-ESS2-5:
• One fall day starts out warm and fairly still. The wind picks up and the temperature drops and it begins to rain.
• The flag outside a school has been resting against the flagpole, unmoving all morning. In the early afternoon, it starts flapping in the wind. At sunset, rain begins.</td></tr>
</table>

	• Fall days were chilly, then the temperature warmed up for a few days. • A tornado formed in the Pacific Ocean near Oregon.

This Performance Expectation and associated Evidence Statements support the following Task Demands.
Task Demands
1. Evaluate the sufficiency and limitations of data collected to explain the phenomenon.
2. Identify the outcome data that should be collected in an investigation of the interactions of air masses and the resulting changes in weather conditions.
3. Make and/or record observations about the interactions of air masses and/or the relationships between those interactions and patterns of weather in a particular location.
4. Describe, illustrate, or select tools, locations, and/or methods to use in investigations of phenomena related to interactions of air masses. This should show how or where measurements will be taken.
5. Identify, select, or describe the relevance of particular data or sources relevant to the process of weather forecasting.
6. Predict the effects of given changes in the air masses' interactions on subsequent weather.
7. Identify or specify inferences supported by data collected.

<table>
<tr><td>Performance Expectation</td><td colspan="3">MS ESS2-6
Develop and use a model to describe how unequal heating and rotation of the Earth cause patterns of atmospheric and oceanic circulation that determine regional climates.</td></tr>
<tr><td>Dimensions</td><td>Developing and Using Models
• Develop and use a model to describe phenomena.</td><td>ESS2.C: The Roles of Water in Earth’s Surface Processes
• Variations in density due to variations in temperature and salinity drive a global pattern of interconnected ocean currents.

ESS2.D: Weather and Climate
• Weather and climate are influenced by interactions involving sunlight, the ocean, the atmosphere, ice, landforms, and living things. These interactions vary with latitude, altitude, and local and regional geography, all of which can affect oceanic and atmospheric flow patterns.
• The ocean exerts a major influence on weather and climate by absorbing energy from the sun, releasing it over time, and globally redistributing it through ocean currents.</td><td>Systems and System Models
• Models can be used to represent systems and their interactions—such as inputs, processes and outputs—and energy, matter, and information flows within systems.</td></tr>
<tr><td>Clarifications and Content Limits</td><td colspan="3">Clarification Statements
• Emphasis is on how patterns vary by latitude, altitude, and geographic land distribution.
• Emphasis of atmospheric circulation is on the sunlight-driven latitudinal banding, the Coriolis Effect, and resulting prevailing winds; emphasis of ocean circulation is on the transfer of heat by the global ocean convection cycle, which is constrained by the Coriolis Effect and the outlines of continents.
• Examples of models can be diagrams, maps and globes, or digital representations.

Content Limits
• Assessment does not include the dynamics of the Coriolis effect.
• <u>Students do not need to know</u>: names of specific winds, different cloud types (cumulus, cirrus etc.), names of specific ocean currents, or perform any quantitative analyses based on the Coriolis Effect, mathematical calculations beyond trends, or measurements of central tendency.</td></tr>
<tr><td>Science Vocabulary Students are Expected to Know</td><td colspan="3">Climate, temperature, atmospheric pressure, density, current, latitude, altitude, Coriolis effect, convection, condensation, precipitation, cloud, water cycle, air mass circulation, vegetation, latitude, longitude, rain shadow.</td></tr>
<tr><td>Science Vocabulary Students are Not Expected to Know</td><td colspan="3">Trade winds, Easterlies, Westerlies, cumulus, cirrus, or other cloud names, Gulf Stream, Labrador, UV rays, horse latitudes, Tropic of Cancer, Tropic of Capricorn.</td></tr>
<tr><td colspan="4" align="center">Phenomena</td></tr>
<tr><td>Context/ Phenomena</td><td colspan="3">Some example phenomena for MS-ESS2-6:</td></tr>
</table>

	• In December 2010, Gary, Indiana, on the southeast shores of Lake Michigan, had approximately 30 inches of snow over a three-day period, whereas Chicago, Illinois, 30 miles away, received barely any snow. • Onshore and offshore breezes—in the morning, the breeze comes in from the ocean. At night, the breeze is blowing in the opposite direction. • Wind storms in the Sahara become hurricanes that affect the east coast of North America and the Caribbean, but not the coast of South America. • The Westerlies vs. The Easterlies and the trade winds—why are these wind patterns banded as you move north from the equator?

This Performance Expectation and associated Evidence Statements support the following Task Demands.
Task Demands
1. Select or identify from a collection of potential model components, including distractors, components needed for a model that can explain the effect of unequal heating of Earth's surface. Components might include oceans, land forms, wind currents, ocean currents, energy flows, upwelling, downwelling, water temperature, air temperature, and salinity.
2. Assemble or complete an illustration or flow chart that is capable of representing the effect of unequal heating of Earth's systems on atmospheric and oceanic circulation. Key components of the model might include: oceans, land forms, wind current, ocean current, energy flows, upwelling, downwelling, water temperature, and salinity.
3. Manipulate the components of a model to demonstrate the changes, properties, processes, and/or events that act to result in a phenomenon.
4. Make predictions about the effects of changes in temperature on a phenomenon. Predictions can be made by manipulating model components, completing illustrations, or selecting from lists with distractors. Make predictions about the effects of changes in water temperature or density, distance from the lake, location, etc.
5. Identify missing components, relationships, or other limitations of a model.
6. Describe, select, or identify the relationships among components of a model that explain the effect of unequal heating of Earth's systems on atmospheric and oceanic circulation.

<table>
<tr><td>Performance Expectation</td><td colspan="3">MS-ESS3-1
Construct a scientific explanation based on evidence for how the uneven distributions of Earth's mineral, energy, and groundwater resources are the result of past and current geoscience processes.</td></tr>
<tr><td>Dimensions</td><td>Constructing Explanations and Designing Solutions
• Construct a scientific explanation based on valid and reliable evidence obtained from sources (including the students' own experiments) and the assumption that theories and laws that describe the natural world operate today as they did in the past and will continue to do so in the future.</td><td>ESS3.A: Natural Resources
• Humans depend on Earth's land, ocean, atmosphere, and biosphere for many different resources. Minerals, fresh water, and biosphere resources are limited, and many are not renewable or replaceable over human lifetimes. These resources are distributed unevenly around the planet as a result of past geologic processes.</td><td>Cause and Effect
• Cause-and-effect relationships may be used to predict phenomena in natural or designed systems.</td></tr>
<tr><td>Clarifications and Content Limits</td><td colspan="3">Clarification Statements
• Emphasis is on how these resources are limited and typically non-renewable, and how their distributions are significantly changing as a result of removal by humans.
• Examples of uneven distributions of resources as a result of past processes include but are not limited to petroleum (location of the burial of organic marine sediments and subsequent geologic traps), metal ores (location of past volcanic and hydrothermal activity associated with subduction zones), and soil (location of active weathering and/or deposition of rock).</td></tr>
<tr><td>Science Vocabulary Students Are Expected to Know</td><td colspan="3">Agricultural, biosphere, conservation, consumption, deposition, distribution, efficient, energy source, geologic trap, hydrothermal, impact, interdependence, marine sediment, metal ore, , organic, petroleum, regulation, renewable energy, subduction zone.</td></tr>
<tr><td>Science Vocabulary Students Are Not Expected to Know</td><td colspan="3">Bitumen, harvesting of resources, viscous, natural gas, oil shale, sustainability, tar sand, extract, irreversible.</td></tr>
<tr><td colspan="4">Phenomena</td></tr>
<tr><td>Context/ Phenomena</td><td colspan="3">Some example phenomena for MS-ESS3-1:
• Large surface deposits of sand and gravel are much more common in Massachusetts than they are in Virginia.
• Diamonds are found on the ground in a State Park in southwestern Arkansas.
• Bauxite, an Aluminum ore, and fossil tree roots are found in an exposure in Queensland, Australia.
• A well is drilled and water is discovered near Colorado Springs, CO. Ten miles to the Southwest, another well is drilled to the same depth and no water is discovered.</td></tr>
<tr><td colspan="4">This Performance Expectation and associated Evidence Statements support the following Task Demands.</td></tr>
<tr><td colspan="4">Task Demands</td></tr>
<tr><td colspan="4">1. Articulate, describe, illustrate, or select the relationships, interactions, and/or processes to be explained. This may entail sorting relevant from irrelevant information or features.</td></tr>
</table>

2. Express or complete a causal chain explaining that the uneven distribution of Earth's mineral, energy, and groundwater resources are the result of past and current geoscience processes. This may include indicating directions of causality in an incomplete model such as a flow chart or diagram, or completing cause-and-effect chains.
3. Identify evidence supporting the inference of causation that is expressed in a causal chain.
4. Use an explanation to predict the distribution of Earth's mineral, energy, or groundwater resources, given a change in current geoscience processes.
5. Describe, identify, and/or select information needed to support an explanation.

<table>
<tr><td>Performance Expectation</td><td colspan="3">MS-ESS3-2
Analyze and interpret data on natural hazards to forecast future catastrophic events and inform the development of technologies to mitigate their effects.</td></tr>
<tr><td>Dimensions</td><td>Analyzing and Interpreting Data
• Analyze and interpret data to determine similarities and differences in findings.</td><td>ESS3.B: Natural Hazards
• Mapping the history of natural hazards in a region, combined with an understanding of related geologic forces, can help forecast the locations and likelihoods of future events.</td><td>Patterns
• Graphs, charts, and images can be used to identify patterns in data.</td></tr>
<tr><td>Clarifications and Content Limits</td><td colspan="3">Clarification Statements
• Emphasis is on how some natural hazards, such as volcanic eruptions and severe weather, are preceded by phenomena that allow for reliable predictions, but others, such as earthquakes, occur suddenly and with no notice, and thus are not yet predictable.
• Examples of natural hazards can be taken from interior processes (such as earthquakes and volcanic eruptions), surface processes (such as mass wasting and tsunamis), or severe weather events (such as hurricanes, tornadoes, and floods).
• Examples of data can include the locations, magnitudes, and frequencies of the natural hazards.
• Examples of technologies can be global (such as satellite systems to monitor hurricanes or forest fires) or local (such as building basements in tornado-prone regions or reservoirs to mitigate droughts).

Content Limits
• Analysis may include recognizing patterns in data, identifying periodicity, straightforward mathematical comparisons (more, less, faster, slower), examining trends, looking for differences in tabular data, qualitative spatial analysis (e.g., looking at fault lines), recognizing trends and patterns. May include drawing lines of best fit and extrapolating from those lines.
• Analysis should not include regression analysis or calculating correlations.</td></tr>
<tr><td>Science Vocabulary Students are Expected to Know</td><td colspan="3">Air mass, air mass circulation, altitude, atmospheric circulation, biosphere, carbon dioxide, climatic pattern, condensation, convection cycle, Coriolis effect, cyclical, density, distribution, geography, geological, gradual, intensity, land distribution, latitude, longitude, ocean circulation, orbit, orientation, pressure, redistribute, salinity, store, tectonic, tectonic cycle, tilt, transfer, unequal heating of air, unequal heating of land masses, unequal heating of oceans, weather map, catastrophic, debris, frequency, geologic, interdependent, magnitude, mass wasting, natural process, reservoir, satellite.</td></tr>
<tr><td>Science Vocabulary Students are Not Expected to Know</td><td colspan="3">Concentration, electromagnetic radiation, radiation, sea level.</td></tr>
<tr><td colspan="4" align="center">Phenomena</td></tr>
<tr><td>Context/ Phenomena</td><td colspan="3">For this performance expectation, the phenomena are sets of data. Those are the observed facts that the kids will look at to discover patterns. Below, we enumerate some of the patterns that might comprise the data sets (phenomena) to be analyzed.

Patterns that describe the data sets for MS-ESS3-2:</td></tr>
</table>

	• A sequence of maps illustrates temperature patterns and occurrence of tornados over the course of the year (to identify variations of tornado risk across regions and also to identify more proximate predictors of tornados). • A sequence of maps illustrates temperature and humidity patterns and occurrence of hurricanes over the course of the year (to identify variations of hurricane risk across regions and also to identify more proximate predictors of hurricanes). • Temperature and humidity patterns in the Pacific Ocean can be correlated to the snow pack on Mt. Hood. • A map of average snowfall in the Great Lakes region shows more snow has fallen in locations nearer to the lakes. Data include surface temperatures, water temperature, wind patterns and snowfall.

This Performance Expectation and associated Evidence Statements support the following Task Demands.
Task Demands
1. Organize/Arrange data to highlight patterns, trends, or correlations between natural hazards and geologic/atmospheric events that occur before a natural hazard.*
2. Tabulate/Graph data to highlight patterns, trends, or correlations between natural hazards and geologic/atmospheric events that occur before a natural hazard.*
3. Use relationships identified in the data to predict natural hazards.
4. Illustrate or describe patterns over time that can be used to predict natural hazards.*
5. Identify human and societal responses designed to mitigate catastrophic natural hazards.

*denotes those task demands which are deemed appropriate for use in stand-alone item development

<table>
<tr><td>Performance Expectation</td><td colspan="3">MS-ESS3-3
Apply scientific principles to design a method for monitoring and minimizing a human impact on the environment.</td></tr>
<tr><td>Dimensions</td><td>Constructing Explanations and Designing Solutions
• Apply scientific principles to design an object, tool, process or system.</td><td>ESS3.C: Human Impacts on Earths Systems
• Human activities have significantly altered the biosphere, sometimes damaging or destroying natural habitats and causing extinction of other species. But changes to Earth’s environments can have different impacts (negative and positive) for different living things.
• Typically as human populations and per-capita consumption of natural resources increase, so do the negative impacts of Earth unless the activities and technologies involved are engineered otherwise.</td><td>Cause and Effect
• Relationships can be classified as causal or correlational, and correlation does not necessarily imply causation.</td></tr>
<tr><td>Clarifications and Content Limits</td><td colspan="3">Clarification Statements
Examples of the design process include examining human environmental impacts, assessing the kinds of solutions that are feasible, and designing and evaluating solutions that could reduce that impact.
Examples of human impacts can include water usage (such as the withdrawal of water from streams and aquifers or the constructions of dams and levees), land usage (such as urban development, agriculture, or the removal of wetlands), and pollution (such as air, water, or land).
Content Limits
Students will not describe the relationship between natural resources and sustainability</td></tr>
<tr><td>Science Vocabulary Students are Expected to Know</td><td colspan="3">Wetland, agriculture, development, fertile, groundwater, industry, material world, mineral, river delta, aquifer, economic, land usage, levee, water usage, consumption, land use management, conservation, preservation</td></tr>
<tr><td>Science Vocabulary Students are Not Expected to Know</td><td colspan="3">Anthropogenic changes, urban development, biomass, degradation, destabilize, geoengineering, ozone, pollutant, sea level, stabilize, waste management, harvesting of resources, cost-benefit</td></tr>
<tr><td colspan="4" align="center">Phenomena</td></tr>
<tr><td>Context/ Phenomena</td><td colspan="3">Engineering performance expectations are built around meaningful design problems rather than phenomena.

Some example design problems for MS-ESS3-3:
Nurdles are small plastic pellets, smaller than a pea. Billions of them are used in creating plastic products. Many fall out of the truck or ship container that they are transported in and end up in oceans where they are mistaken as food by marine animals.
Glen Canyon Dam is located on the Arizona and behind it sits Lake Powell the second largest reservoir in the United States. Glen Canyon Dam holds back sediment that would naturally replenish downstream ecosystems. The sediment that is trapped behind the dam is filling Lake Powell at a rate of roughly 100 million tons of sediment a year, decreasing the dam’s ability to store water.</td></tr>
</table>

	• Farmers in Iowa plow their fields in the spring in order to break up the thick soil and disrupt weeds from growing. The practice of plowing however, causes farmers to lose valuable top soil due to wind erosion. • In the central North Pacific Ocean there is what is described as a great garbage patch. This large area has high concentrations of plastics, fishing nets, and other debris. This debris is sometimes mistaken as food by marine animals.
This Performance Expectation and associated Evidence Statements support the following Task Demands.	
Task Demands	
1. Identify or assemble from a collection, including distractors, the relevant aspects of human impact on the environment that given design solutions, if implemented, will resolve/improve.	
2. Using the given information about human impact on the environment, select or identify the criteria against which the device or solution should be judged.	
3. Using given information about human impact on the environment, select or identify constraints that the device or solution must meet.	
4. Using given data, propose/illustrate/assemble a potential device (prototype) or solution to monitor and/or minimize human impact on the environment.	
5. Using a simulator, test a proposed prototype and evaluate the outcomes, potentially including proposing and testing modifications to the prototype.	

<table>
<tr><td>Performance Expectation</td><td colspan="3">MS-ESS3-4
Construct an argument supported by evidence for how increases in human population and per-capita consumption of natural resources impact Earth’s systems.</td></tr>
<tr><td>Dimensions</td><td>Engaging in Argument from Evidence
• Construct an oral and written argument supported by empirical evidence and scientific reasoning to support or refute an explanation or a model for a phenomenon or a solution to a problem.</td><td>ESS3.C: Human Impacts on Earth Systems
• Typically, as human populations and per-capita consumption of natural resources increase, so do the negative impacts on Earth, unless the activities and technologies involved are engineered otherwise.</td><td>Cause and Effect
• Cause and effect relationships may be used to predict phenomena in natural or designed systems.</td></tr>
<tr><td>Clarifications and Content Limits</td><td colspan="3">Clarification Statements
• Examples of evidence include grade-appropriate databases on human populations and the rates of consumption of food and natural resources such as fresh water, minerals, and energy sources.
• Examples of impacts can include changes to the appearance, composition, and structure of Earth’s systems as well as the rates at which they change. The consequences of increases in human populations and consumption of natural resources are described by science, but science does not make the decisions for the actions society takes.

Content Limits
• Assessment is limited to one form of consumption and its associated impacts.
• <u>Students do not need to know</u>: mechanisms or details about interior geological processes, quantities and types of pollution released, changes to biomass and species diversity, or changes in land surface use.</td></tr>
<tr><td>Science Vocabulary Students Are Expected to Know</td><td colspan="3">Conservation, recycling, perishable, synthetic, manufactured, rivers, lakes, groundwater, fertile, delta, fossil fuels, pollution, composition, glacier, mass, volume, concentration.</td></tr>
<tr><td>Science Vocabulary Students Are Not Expected to Know</td><td colspan="3">Tar sands, oil shales, agricultural efficiency, urban planning, aesthetics, biomass, glacial ice volumes, hydrosphere, cryosphere, geosphere, acidification, empirical evidence, polar caps.</td></tr>
<tr><td colspan="4">Phenomena</td></tr>
<tr><td>Context/ Phenomena</td><td colspan="3">Some example phenomena for MS-ESS3-4:
• Lake Urmia in Iran was once the nation’s largest lake. Today, the lake is 5% as large as it used to be.
• In 1990, much of the tropical rain forests on the Hainan Island were clear-cut to obtain wood, and to create space for plantations. Today, the forests are still smaller and less developed than they were before 1990.
• A coal power plant in Martins Lake, Texas, releases huge clouds of gas into the air every day.
• The open-pit copper mine Ok Tedi Mine in Papua, New Guinea, releases its drainage nearby. Downstream, the rivers turned orange and the fish died.</td></tr>
</table>

This Performance Expectation and associated Evidence Statements support the following Task Demands.
Task Demands
1. Articulate, describe, illustrate, or select the relationships, interactions, and/or processes to be explained. This may entail sorting relevant from irrelevant information.
2. Predict outcomes when properties or amounts of consumption are changed, given the inferred cause and effect relationships.
3. Describe, identify, and/or select information needed to support an explanation of how increases in human population and per-capita consumption of natural resources impact Earth's systems.
4. Identify patterns or evidence in the data that support conclusions about the relationship between per capita consumption and limited natural resources.*
5. Using evidence, explain the relationship between per capita consumption and limited natural resources.*
6. Manipulate the components of a model to demonstrate the changes, properties, processes, and/or events that act to result in the phenomenon.*

*denotes those task demands which are deemed appropriate for use in stand-alone item development

<table>
<tr><td>Performance Expectation</td><td colspan="3">MS-ESS3-5
Ask questions to clarify evidence of the factors that have caused the rise in global temperatures over the past century.</td></tr>
<tr><td>Dimensions</td><td>Asking Questions and Defining Problems
• Ask questions to identify and clarify evidence of an argument.</td><td>ESS3.D: Global Climate Change
• Human activities, such as the release of greenhouse gases from burning fossil fuels, are major factors in the current rise in Earth’s mean surface temperature (global warming). Reducing the level of climate change and reducing human vulnerability to whatever climate changes do occur depend on the understanding of climate science, engineering capabilities, and other kinds of knowledge, such as the understanding of human behavior and on applying that knowledge wisely in decisions and activities.</td><td>Stability and Change
• Stability might be disturbed either by sudden events or gradual changes that accumulate over time.</td></tr>
<tr><td>Clarifications and Content Limits</td><td colspan="3">Clarification Statements
• Examples of factors include human activities (such as fossil-fuel combustion, cement production, and agricultural activity) and natural processes (such as changes in incoming solar radiation or volcanic activity).
• Examples of evidence can include tables, graphs, and maps of global and regional temperatures; atmospheric levels of gases such as carbon dioxide and methane; and the rates of human activities. Emphasis is on the major role that human activities play in causing the rise in global temperatures.</td></tr>
<tr><td>Science Vocabulary Students Are Expected to Know</td><td colspan="3">Force, rotation, intensity, physical change, glacial, weather condition, natural resource, natural process, catastrophic, cycle, atmospheric composition, environmental, pollution, societal, renewable resource, nonrenewable, oil, absorb.</td></tr>
<tr><td>Science Vocabulary Students Are Not Expected to Know</td><td colspan="3">Climactic pattern, cyclical, concentration, magnitude, destabilize, consumption, civilization, degradation, pollutant, sea level, stable, natural gas.</td></tr>
<tr><td colspan="4">Phenomena</td></tr>
<tr><td>Context/ Phenomena</td><td colspan="3">Some example phenomena for MS-ESS3-5:
• A region in the Saint Elias Mountains in Alaska used to be covered by Plateau Glacier. It is now populated with thick vegetation and lake.
• On December 14th, 2016, the Deely Power Plant was operating. Its chimney emitted a large cloud of white smoke.
• The Solomon Islands are a group of small islands located in the Pacific Ocean. Five of these islands disappeared in 2016.
• Mount Etna, one of the world’s most active volcanoes, erupted in May 2016, delivering large plumes of smoke that filled the horizon.</td></tr>
<tr><td colspan="4">This Performance Expectation and associated Evidence Statements support the following Task Demands.</td></tr>
<tr><td colspan="4">Task Demands</td></tr>
<tr><td colspan="4">6. Organize and/or arrange (e.g., using illustrations and/or labels) or summarize data to highlight trends, patterns, or correlations.</td></tr>
<tr><td colspan="4">7. Generate/construct graphs, tables, or assemblages of illustrations and/or labels of data that document patterns, trends, or correlations relating to climate change. This may include sorting out distractors.</td></tr>
</table>

8. Express or complete a causal chain explaining the effects that climate change has on the environment. This may include indicating directions of causality in an incomplete model such as a flow chart or diagram or completing cause-and-effect chains.
9. Compile, from given information, the particular data needed for a particular inference about the relationship between greenhouse gas emissions and rising global temperatures. This can include sorting out the relevant data from the given information.
10. Describe, select, or identify the relationships among components of a model that describe the mechanism of rising global temperatures, or explain the consequences of rising global temperatures.
11. Select, from a list of potential hypotheses including distractors, either the testable hypothesis from untestable hypotheses or the best hypothesis to clarify evidence relating to climate change.
12. Construct or assemble a valid hypothesis that clarifies evidence relating to climate change.
13. Select from a list of questions, including distractors, about the relationships among the data that either support or contradict a hypothesis or to clarify data that describe the mechanism of rising global temperatures, or explain the consequences of rising global temperatures.
14. Ask questions to obtain or clarify information related to the rise of global temperatures in the past century.

Specifications for High School

Introduction

This document presents *cluster specifications* for use with the Next Generation Science Standards (NGSS). These standards are based on the Framework for K-12 Science Education. The present document is not intended to replace the standards, but rather to present guidelines for the development of items and item clusters used to measure those standards.

The remainder of this section provides a very brief introduction to the standards and the framework, an overview of the design and intent of the item clusters, and a description of the cluster specifications that follow. The bulk of the document is composed of cluster specifications, organized by grade and standard.

Background on the framework and standards

The Framework for K-12 Science Education are organized around three core dimensions of scientific understanding. The standards are derived from these same dimensions:

- Disciplinary Core Ideas: The fundamental ideas that are necessary for understanding a given science discipline. The core ideas all have broad importance within or across science or engineering disciplines, provide a key tool for understanding or investigating complex ideas and solving problems, relate to societal or personal concerns, and can be taught over multiple grade levels at progressive levels of depth and complexity.

- Science and Engineering Practices: The practices are what students DO to make sense of phenomena. They are both a set of skills and a set of knowledge to be internalized. The SEPs reflect the major practices that scientists and engineers use to investigate the world and design and build systems.

- Cross-Cutting Concepts: These are concepts that hold true across the natural and engineered world. Students can use them to make connections across seemingly disparate disciplines or situations, connect new learning to prior experiences, and more deeply engage with material across the other dimensions. The NGSS requires that students explicitly use their understanding of the CCCs to make sense of phenomena or solve problems.

- There is substantial overlap between and among the three dimensions. For example, the cross-cutting concepts are echoed in many of the disciplinary core ideas. The core ideas are often closely intertwined with the practices. This overlap reflects the nature of science itself. For example, we often come to understand and communicate causal relationships by employing models to make sense of observations. Even within a dimension, overlap exists. Quantifying characteristics of phenomena is important in developing an understanding of them, so employing computational and mathematical thinking in the construction and use of models is a very common scientific practice, and one of the cross-cutting concepts suggests that scientists often infer causality by observing patterns. In short, the dimensions are not orthogonal.

The framework envisions effective science education as occurring at the intersection of these interwoven dimensions: students learn science by doing science—applying the practices through the lens of the cross-cutting concepts to investigate phenomena that relate to the content of the disciplinary core ideas.

Item clusters

Each item cluster is designed to engage the examinee in a grade-appropriate, meaningful scientific activity aligned to a specific standard.

Each cluster begins with a phenomenon, an observable fact or design problem that engages student interest and can be explained, modeled, investigated, or designed using the knowledge and skill described by the standard in question.

What it means to be observable varies across practices. For example, a phenomenon for a performance expectation exercising the a*nalyze data* practice may be observable through regularities in a data set, while standards related to the *development and use of models* might be something that can be watched, seen, felt, smelled, or heard.

What it means to be observable also varies across grade levels. For example, elementary-level phenomena are very concrete and directly observable. At the high school level, an observation of the natural world may be more abstract--for example, "observing" changes in the chemical composition of cells through the observation of macroscopic results of those changes on organism physiology, or through the measurement of system- or organ-level indications.

Content limits refine the intent of the performance expectations and provide limits on what may be asked of items in the cluster to structure the student activity. The content limits also reflect the disciplinary core ideas learning progressions that are present in the K-12 Framework for Science Education.

The task or goal should be explicitly stated in the stimulus or the first item in the cluster: statements such as "In the questions that follow, you will develop a model that will allow you to identify moons of Jupiter," or "In the questions below, you will complete a model to describe the processes that lead to the steam coming out of the teapot."

Whereas item clusters have been described elsewhere as "scaffolded," they are better described as providing structure to the task. For example, some clusters begin with students summarizing data to discover patterns that may have explanatory value. Depending on the grade level and nature of the standard, items may provide complete table shells or labeled graphs to be drawn, or may require the student to choose what to tabulate or graph. Subsequent items may ask the student to note patterns in the tabulated or graphed data and draw on domain content knowledge to posit explanations for the patterns.

These guidelines for clusters do not appear separately in the specifications. Rather, they apply to all clusters.

Structure of the cluster specifications

The item cluster specifications are designed to guide the work of item writers and the review of item clusters by stakeholders.

Each item cluster has the following elements:

- The text of the performance expectations, including the practice, core idea, and cross-cutting concept.
- Content limits, which refine the intent of the performance expectations and provide limits of what may be asked of examinees. For example, they may identify the specific formulae that students are expected to know or not know.
- Vocabulary, which identifies the relevant technical words that students are expected to know, and related words that they are explicitly not expected to know. The vocabulary included in both sections (Students are Expected to Know and Students Are Not Expected to Know) were developed after the reviews of standards at the current/preceding grades, the original NGSS documentation, and item writer reference documentation including the Children Writers' Word Book and ED Core Vocabularies in Reading, Mathematics, Science and Social Studies. All vocabulary included in the specifications was reviewed and edited by teacher committees during the specification reviews by states. Of course, the latter category should not be considered exhaustive, since the boundaries of relevance are ambiguous, and the list is limited by the imagination of the writers/reviewers.
- Sample phenomena, which provide some examples of the sort of phenomena that would support effective item clusters related to the standard in question. In general, these should be guideposts, and item writers should seek comparable phenomena, rather than drawing on those within the documents. Novelty is valued when applying scientific practices.
- Task demands comprise the heart of the specifications. These statements identify the types of items and activities that item writers should use, and each item written should be clearly linked to one or more of the demands. The verbs in the demands (e.g., *select, identify, illustrate, describe*) provide guidance on the types of interactions that item writers might employ to elicit the student response. We avoid explicitly identifying interaction types or item formats to accommodate future innovations and to avoid discouraging imaginative

work by the item writers. All task demands should be aligned to a minimum of one of the three dimensions (DCI, SEP and CCCs) and across task demands within a cluster, all three dimensions must be addressed.

- For each cluster we present, the printed documentation includes the cluster, the task demands represented by each item, and its linkage to the content, practice and cross-cutting concept identified in the performance expectation.

Item cluster specifications follow, organized by domain and standard.

Performance Expectation	**HS-PS1-1** Use the periodic table as a model to predict the relative properties of elements based on the patterns of electrons in the outermost energy level of atoms.		
Dimensions	**Developing and Using Models** • Use a model to predict the relationships between systems or between components of a system.	**PS1.A: Structure and Properties of Matter** • Each atom has a charged substructure consisting of a nucleus, which is made of protons and neutrons, surrounded by electrons. • The periodic table orders elements horizontally by the number of protons in the atom's nucleus and places those with similar chemical properties in columns. The repeating patterns of this table reflect patterns of outer electron states.	**Patterns** • Different patterns may be observed at each of the scales at which a system is studied and can provide evidence for causality in explanations of phenomena.
Clarifications and Content Limits	**Clarification Statements** • Examples of properties that could be predicted from patterns could include reactivity of metals, types of bonds formed, numbers of bonds formed, and reactions with oxygen. **Content Limits** • Assessment is limited to main group elements. • Assessment does not include quantitative understanding of ionization energy beyond relative trends. • <u>Students do not need to know:</u> Properties of individual elements, names of groups, anomalous electron configurations (Chromium and Copper)		
Science Vocabulary Students are Expected to Know	Proton, electron, neutron, valence shell, filled shell, ion, cation, anion, metal, nonmetal, metalloid, group, period, family, pure substance, atomic number, atomic symbol, atomic weight, ionic bond, covalent bond, s, p, d, f orbitals, electron configuration, core electrons, nucleus, single, double, triple bond(s), molar mass, atomic radius, electronegativity,		
Science Vocabulary Students are Not Expected to Know	Oxidation state, diatomic, polyatomic ions, empirical formulas, molecular formulas, quantum, photon, Heisenberg Uncertainty Principle, Hund's Rule, Pauli Exclusion Principle		
Phenomena			
Context/ Phenomena	Some example phenomena for HS-PS1-1: • Potassium chloride (KCl) tastes similar to table salt (sodium chloride (NaCl)). • Balloons are filled with helium gas instead of hydrogen gas. • Scientists work with silicate substrates in chambers filled with Argon instead of air. • Diamond, graphene, and fullerene are different molecules/materials that are only made of carbon.		
This Performance Expectation and associated Evidence Statements support the following Task Demands.			
Task Demands			
1. Select or identify from a collection of periodic table components (periods, groups, etc.), including distractors, the components needed to model the phenomenon.			
2. Make predictions about the properties of elements based on the number of valence electrons. Predictions can be made by completing illustrations or selecting from lists with distractors.			

3. Identify missing components, relationships, or other limitations of the model. (Hydrogen similar to Alkali metals, one valence electron, and Halogens, missing only one valence electron).
4. Describe, select, or identify the relationships among components of the periodic table that describe the properties of valence electrons, or explains the properties of elements.

Performance Expectation	**HS-PS1-2** Construct and revise an explanation for the outcome of a simple chemical reaction based on the outermost electron states of atoms, trends in the periodic table, and knowledge of the patterns of chemical properties.		
Dimensions	**Constructing explanations and designing solutions** • Construct and revise an explanation based on valid and reliable evidence obtained from a variety of sources (including students' own investigations, models, theories, simulations, peer review) and the assumption that theories and laws that describe the natural world operate today as they did in the past and will continue to do so in the future.	**PS1.A: Structure and Properties of Matter** • The periodic table orders elements horizontally by the number of protons in the atom's nucleus and places those with similar chemical properties in columns. The repeating patterns of this table reflect patterns of outer electron states. **PS1.B: Chemical Reactions** • The fact that atoms are conserved, together with knowledge of the chemical properties of the elements involved, can be used to describe and predict chemical reactions.	**Patterns** • Different patterns may be observed at each of the scales at which a system is studied and can provide evidence for causality in explanations of phenomena.
Clarifications and Content Limits	**Clarification Statements** • Examples of chemical reactions could include the reaction of sodium and chlorine, of carbon and oxygen, or of carbon and hydrogen. **Content Limits** • Assessment is limited to chemical reactions involving main group elements and combustion reactions.		
Science Vocabulary Students are Expected to Know	Reversible, atomic weight, chemical bond, electron sharing, ion, outer electron state, energy level, electron transfer, concentration, equilibrium, endothermic, exothermic, stable, combustion, yield(s), flammability, octet		
Science Vocabulary Students are Not Expected to Know	Molecular orbital diagram, multiplicity, antibonding orbitals, rearrangement, by-product, oxidation-reduction reaction, decomposition, single replacement reaction, double replacement reaction, synthesis reaction, precipitate		
Phenomena			
Context/ Phenomena	Some example phenomena for HS-PS1-2: • A coal oven without proper ventilation produces billows of dark smoke. • Two metals are placed in water. One bubbles and fizzes, while the other gives off a yellow flame and white smoke. • Carlsbad Caverns is a large cave in New Mexico. Inside, large pointy structures appear to be growing from the ceiling. • A shiny metallic solid is combined with a green gas, resulting in a white crystalline solid.		
This Performance Expectation and associated Evidence Statements support the following Task Demands.			
Task Demands			

15. Use relationships identified in the data to predict properties of other chemical compounds, elements, and/or mixtures.
16. Identify patterns or evidence in the data that supports inferences about the properties of other chemical compounds/elements/mixtures.
17. Organize and/or arrange (e.g., using illustrations and/or labels), or summarize data to highlight trends, patterns, or correlations.
18. Articulate, describe, illustrate, or select the relationships, interactions, and/or processes to be explained. This may entail sorting relevant from irrelevant information or features.
19. Use an explanation to predict the properties of other chemical compounds/elements/mixtures given a change in reagents or conditions.
20. Generate/construct graphs, tables, or assemblages of illustrations and/or labels of data that document patterns, trends, or correlations relating to the periodic table. This may include sorting out distractors.
21. Select, articulate, or construct an explanation about a chemical reaction. This may include identifying/selecting the products of the reaction as part of an explanation.

<table>
<tr><td>Performance Expectation</td><td colspan="3">HS-PS1-3
Plan and conduct an investigation to gather evidence to compare the structure of substances at the bulk scale to infer the strength of electrical forces between particles.</td></tr>
<tr><td>Dimensions</td><td>Planning and Carrying Out Investigations
• Plan and conduct an investigation individually and collaboratively to produce data to serve as the basis for evidence, and in the design: decide on types, how much, and accuracy of data needed to produce reliable measurements and consider limitations on the precision of the data (e.g., number of trials, cost, risk, time), and refine the design accordingly.</td><td>PS1.A Structure and Properties of Matter
• The structure and interactions of matter at the bulk scale are determined by electrical forces within and between atoms.</td><td>Patterns
• Different patterns may be observed at each of the scales at which a system is studied and can provide evidence for causality in explanations of phenomena.</td></tr>
<tr><td>Clarifications and Content Limits</td><td colspan="3">Clarification Statements
• Emphasis is on understanding the strength of forces between particles, not on naming specific intermolecular forces (such as dipole-dipole).
• Examples of particles could include ions, atoms, molecules, and networked materials (such as graphite).
• Examples of bulk properties of substances could include the melting point and boiling point, vapor pressure, and surface tension.

Content Limits
• Assessment does not include Raoult’s law, nor calculations of vapor pressure.</td></tr>
<tr><td>Science Vocabulary Students Are Expected to Know</td><td colspan="3">Nucleus, proton, electron, neutron, electron cloud, intramolecular force, covalent bond, ionic bond, intermolecular force, electrostatic force, electronegativity, electron distribution, polarity, temporary polarity, permanent polarity, polarize, surface area, atomic radius, atomic weight, atomic mass, solute, solvent, dissolve.</td></tr>
<tr><td>Science Vocabulary Students Are Not Expected to Know</td><td colspan="3">Dipole, induced dipole, dipole moment, delta, Coulomb’s law, dipole-dipole, London forces, Van der Waals forces, ion-dipole, hydrogen bonding, pi-electron cloud, pi stacking, colligative properties, electron shielding.</td></tr>
<tr><td colspan="4">Phenomena</td></tr>
<tr><td>Context/ Phenomena</td><td colspan="3">Some example phenomena for HS-PS1-3:
• Two neighbors apply different salt treatments to their driveways the night before a freeze is predicted. The next morning, no ice formed on one of their driveways. However, the other driveway was covered with a thin layer of ice.
• A chef makes salad dressing by completely mixing oil, water, and vinegar in a large container. Afterwards, he pours the mixed dressing from the large container into individual containers and places one container on each of the restaurant’s tables before leaving for the night. In the morning, the chef finds a layer of oil floating on top of a liquid layer in each of the containers on the tables.
• After working with painting oils, an artist finds that she must wash her hands with soap and water to remove the oil from her hands, as rinsing with water alone does not remove the oil.</td></tr>
</table>

	• A glass is completely filled with water. When coins are added to the full glass of water, the surface of the water rises above the rim of the glass without spilling.

This Performance Expectation and associated Evidence Statements support the following Task Demands.
Task Demands
1. Identify from a list, including distractors, the materials/tools needed for an investigation of the physical properties/interactions of atomic and/or molecular substances at the bulk scale to gather evidence about the strengths of the electrostatic attractions between the particles of those substances.
2. Identify the outcome data that should be collected in an investigation of the physical properties/interactions of atomic and/or molecular substances at the bulk scale to gather evidence about the strengths of the electrostatic attractions between the particles of those substances.
3. Evaluate the sufficiency and limitations of collected data about the physical properties/interactions of substances at the bulk scale to explain the phenomenon.
4. Make and/or record observations about the physical properties/interactions of substances at the bulk scale that provide evidence to support inferences about the relative strengths of the electrostatic attractions between the particles of those substances.
5. Interpret, summarize, and/or communicate the data from an investigation concerning the physical properties/interactions of substances at the bulk scale.
6. Explain or describe the causal processes that lead to the observed data.
7. Select, describe, or illustrate a prediction concerning the physical properties of or interactions between additional substance(s), and/or the strength of electrostatic attractions between the particles of additional substance(s), made by applying the findings from an investigation.

Performance Expectation	**HS-PS1-4** Develop a model to illustrate that the release or absorption of energy from a chemical reaction system depends upon the changes in total bond energy.		
Dimensions	**Developing and Using Models** • Develop a model based on evidence to illustrate the relationships between systems or between components of a system.	**PS1.A: Structure and Properties of Matter** • A stable molecule has less energy than the same set of atoms separated; one must provide at least this energy in order to take the molecule apart. **PS1.B: Chemical Reactions** • Chemical processes, their rates, and whether or not energy is stored or released can be understood in terms of the collisions of molecules and the rearrangements of atoms into new molecules, with consequent changes in the sum of all bond energies in the set of molecules that are matched by changes in kinetic energy.	**Energy and Matter** • Changes of energy and matter in a system can be described in terms of energy and matter flows into, out of, and within that system.
Clarifications and Content Limits	**Clarification Statements** • Emphasis is on the idea that a chemical reaction is a system that affects the energy change. Examples of models could include molecular-level drawing and diagrams of reactions, graphs showing the relative energies of reactants and products, and representations showing energy is conserved. **Content Limits** • *Assessment does not include calculating the total bond energy changes during a chemical reaction from the bond energies of reactants and products.*		
Science Vocabulary Students are Expected to Know	Transfer, heat energy, atomic arrangement, stored energy, conversion, bond energy, release of energy, endothermic, exothermic		
Science Vocabulary Students are Not Expected to Know	Recombination of chemical elements, stable, chemical system, chemical reaction rate		
Phenomena			
Context/ Phenomena	Some example phenomena for HS-PS1-4: • Scientists gather samples of rock from the ocean floor. One sample looks and feels like ice, but burns and produces a flame when ignited. • Wet cement is left sitting outside. After one day, the cement becomes hard and stiff. • A temperature of a sample of tin is lowered from room temperature to 0 °C. The tin changes color from silver to gray, becomes brittle, and starts developing cracks on its surface. • Baking soda is added to a container of citric acid at room temperature. The resulting solution becomes cold, and returns back to room temperature after 2 minutes.		
This Performance Expectation and associated Evidence Statements support the following Task Demands.			
Task Demands			

1. Select or identify from a collection of potential model components, including distractors, the components needed to model the phenomenon. Components might include models of bonds breaking and forming, heat absorbed or released, or aspects of a chemical reaction.
2. Assemble or complete, from a collection of potential model components, an illustration or flow chart that is capable of representing a release or absorption of energy from a chemical reaction. This does not include labeling an existing diagram.
3. Manipulate the components of a model to demonstrate the changes, properties, processes, and/or events that act to result in the phenomenon.
4. Make predictions about the effects of changes in bond energies. Predictions can be made by manipulating model components, completing illustrations, or selecting from lists with distractors.
5. Describe, select, or identify the relationships among components of a model that describes a release or absorption in energy, or explains why a release or absorption in energy is dependent on total bond energy.

<table>
<tr><td>Performance Expectation</td><td colspan="3">HS-PS1-5
Apply scientific principles and evidence to provide an explanation about the effects of changing the temperature or concentration of the reacting particles on the rate at which a reaction occurs.</td></tr>
<tr><td>Dimensions</td><td>Constructing Explanations and Designing Solutions
• Apply scientific principles and evidence to provide an explanation of phenomena and solve design problems, taking into account possible unanticipated effects.</td><td>PS1.B: Chemical Reactions
• Chemical processes, their rates, and whether or not energy is stored or released can be understood in terms of the collisions of molecules and the rearrangements of atoms into new molecules, with consequent changes in the sum of all bond energies in the set of molecules that are matched by changes in kinetic energy.</td><td>Patterns
• Different patterns may be observed at each of the scales at which a system is studied and can provide evidence for causality in explanations of phenomena.</td></tr>
<tr><td>Clarifications and Content Limits</td><td colspan="3">Clarification Statements
• Emphasis is on student reasoning that focuses on the number and energy of collisions between molecules.

Content Limits
• Assessment is limited to simple reactions in which there are only two reactants; evidence from temperature, concentration, and rate data; and qualitative relationships between rate and temperature.</td></tr>
<tr><td>Science Vocabulary Students are Expected to Know</td><td colspan="3">Stored energy, heat energy, atomic arrangement, conversion, bond energy, endothermic, exothermic, concentration, reaction rate, activation energy, catalyst, enzyme, equilibrium</td></tr>
<tr><td>Science Vocabulary Students are Not Expected to Know</td><td colspan="3">Recombination of chemical elements, stable, chemical system, rate laws, Le Chatelier’s principle, rate constant, zero order reactions, first order reactions, stepwise reactions, rate-determining step, steady state, half-life, free radicals, entropy, Gibb’s free energy</td></tr>
<tr><td colspan="4">Phenomena</td></tr>
<tr><td>Context/ Phenomena</td><td colspan="3">Some example phenomena for HS-PS1-5:
• One bowl of bread dough was set aside to rise in a cool area of a kitchen. Another was set aside to rise near the warm oven. The dough near the oven rose faster than the dough set in the cool area.
• A marble stone was exposed to rain water with different acidities on two different spots on the stone. After some time, one spot on the stone was more eroded than the other.
• Cookies baked in an oven set to 170℃ took longer to bake than cookies baked in an oven set to 220℃.
• Inside a fume hood, an adult wearing gloves and goggles carefully added hydrochloric acid to a solution containing sodium thiosulfate ($Na_2S_2O_3$) and a yellow solid appeared in the test tube. Then, dilute hydrochloric acid was added to a second test tube of $Na_2S_2O_3$, and the yellow solid took longer to appear.</td></tr>
<tr><td colspan="4">This Performance Expectation and associated Evidence Statements support the following Task Demands.</td></tr>
<tr><td colspan="4">Task Demands</td></tr>
</table>

1. Articulate, describe, illustrate, or select the relationships, interactions, and/or processes to be explained. This may entail sorting relevant from irrelevant information or features.
2. Express or complete a causal chain explaining how temperature and/or concentration changes can change the rate of a chemical reaction. This may include indicating directions of causality in an incomplete model such as a flow chart or diagram, or completing cause and effect chains.*
3. Identify patterns or evidence in the data that supports inferences about the effects of changing temperature or concentration on the rate at which a chemical reaction occurs.
4. Use an explanation to predict the changes in the rate of other chemical reactions, given a change in reagents or conditions, including temperature and concentration of reactants.
5. Select, articulate, or construct an explanation about a chemical reaction. This may include identifying/selecting the products of the reaction as part of an explanation.*
6. Use evidence to construct an explanation of how changing temperature or concentration of reacting particles on the rate of a reaction.*

*denotes task demands that are approved for use with standalones.

<table>
<tr><td>Performance Expectation</td><td colspan="3">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.</td></tr>
<tr><td>Dimensions</td><td>Constructing Explanations and Designing Solutions
• Refine a solution to a complex real-world problem, based on scientific knowledge, student-generated sources of evidence, prioritized criteria, and tradeoff considerations.</td><td>PS1.B: Chemical Reactions
• In many situations, a dynamic and condition-dependent balance between a reaction and the reverse reaction determines the numbers of all types of molecules present.

ETS1.C: Optimizing the Design Solution
• Criteria may need to be broken down into simpler ones that can be approached systematically, and decisions about the priority of certain criteria over others (tradeoffs) may be needed. (secondary)</td><td>Stability and Change
• Much of science deals with constructing explanations of how things change and how they remain stable.</td></tr>
<tr><td>Clarifications and Content Limits</td><td colspan="3">Clarification Statements
• Emphasis is on the application of Le Chatelier’s Principle and on refining designs of chemical reaction systems, including descriptions of the connection between changes made at the macroscopic level and what happens at the molecular level.
• Examples of designs could include different ways to increase product formation including adding reactants or removing products

Content Limits
• Assessment is limited to specifying the change in only one variable at a time. Assessment does not include calculating equilibrium constants and concentrations.</td></tr>
<tr><td>Science Vocabulary Students are Expected to Know</td><td colspan="3">Surface area of reactants, dynamic, thermal energy, heat energy, atomic arrangement, equilibrium, bond energy, endothermic, exothermic, catalyst, chemical bond, mole, element, compound, concentration, Le Chatelier’s principle</td></tr>
<tr><td>Science Vocabulary Students are Not Expected to Know</td><td colspan="3">Recombination of chemical elements, stable, chemical system, chemical reaction rate</td></tr>
<tr><td colspan="4" align="center">Phenomena</td></tr>
<tr><td>Context/ Phenomena</td><td colspan="3">Some example phenomena for HS-PS1-6:
• Leftover food left on the counter overnight spoils more quickly than food stored in the refrigerator at 1.6°C.
• Several drops of hydrochloric acid were added to an orange mixture of water and potassium dichromate (K_2CrO_7). The mixture turned yellow.
• In the 1970s scientists observed that the concentration of ozone (O_3) in the upper atmosphere began decreasing.
• A bottle of carbonated soda appears to have fewer bubbles before it is opened than after it is opened.</td></tr>
<tr><td colspan="4" align="center">This Performance Expectation and associated Evidence Statements support the following Task Demands.</td></tr>
</table>

Task Demands
9. Express or complete a causal chain explaining the chemical processes that resulted in a shift in equilibrium. This may include indicating directions of causality in an incomplete model such as a flow chart or diagram, or completing cause and effect chains.
10. Describe, identify, and/or select evidence supporting the inference of causation that is expressed in a causal chain and/or an explanation of the processes that cause the observed results.
11. Predict the direction or the relative magnitude of a change in equilibrium of a chemical system, given a change in the amount/concentration of chemical substances in the system, the temperature of the substances in the system, and/or the amount of pressure applied to the substances in a system.
12. Identify or assemble from a collection, including distractors, of the relevant aspects of the problem that a given design solution, if implemented, will resolve or improve.
13. Using the given information, select or identify the criteria against which the solution should be judged.
14. Using given data, propose, illustrate, or assemble a potential solution that would shift equilibrium to favor the products of a chemical reaction.
15. Using a simulator, test a proposed solution and evaluate the outcomes, potentially including proposing and testing modifications to the solution.

Performance Expectation	**HS-PS1-7** Use mathematical representations to support the claim that atoms, and therefore mass, are conserved during a chemical reaction.		
Dimensions	**Using Mathematics and Computational Thinking** • Use mathematical representations of phenomena to support claims.	**PS1.B Chemical Reactions** • The fact that atoms are conserved, together with knowledge of the chemical properties of the elements involved, can be used to describe and predict chemical reactions.	**Energy and Matter** • The total amount of energy and matter in closed systems is conserved.
Clarifications and Content Limits	**Clarification Statements** • Emphasis is on using mathematical ideas to communicate the proportional relationships between masses of atoms in the reactants and the products, and the translation of these relationships to the macroscopic scale using the mole as the conversion from the atomic to the macroscopic scale. • Emphasis is on assessing students' use of mathematical thinking and not on memorization and rote application of problem-solving techniques. **Content Limits** • Assessment does not include complex chemical reactions. • <u>Students do not need to know:</u> Properties of individual elements		
Science Vocabulary Students are Expected to Know	mole, molar ratio, molar mass, limiting reactant, excess reactant, yield(s), theoretical yield, actual yield, concentration, conversion, reversible, ion, cation, anion, metal, nonmetal, metalloid, pure substance, atomic number, atomic symbol, atomic weight, ionic bond, covalent bond		
Science Vocabulary Students are Not Expected to Know	Dimensional analysis, stoichiometry, (dynamic) equilibrium, La Chatlier's Principle, oxidation state, diatomic, polyatomic ion, empirical formula, by-product, oxidation-reduction reaction, decomposition, single replacement reaction, double replacement reaction, synthesis reaction, combustion reaction, precipitate, solvent, solute, reaction rate, recombination of chemical elements, stable		
Phenomena			
Context/ Phenomena	Some example phenomena for HS-PS1-7: • Methane gas flows into a Bunsen burner. When a spark is applied, methane gas reacts with oxygen in the air to produce a blue flame. The flame gets larger as the oxygen valve is turned to allow more oxygen to mix with methane. • Different masses of baking soda are placed inside three balloons of the same size. Three grams of baking soda is added to the first balloon, four grams is added to the second balloon, and five grams is added to the third balloon. Each balloon is placed on top of a bottle containing 200mL of vinegar, with care that no baking soda is lost from the balloons. When the baking soda inside each balloon drops into the vinegar, the balloons eventually inflate. The balloon containing 4g of baking soda inflates to a larger size than the balloon containing 3g. However, the balloon containing 5g of baking soda inflates to the same size as the balloon containing 4g. • When colorless solutions of sodium sulfate (Na_2SO_4) and strontium nitrate ($Sr(NO_3)$2) are mixed, a white solid forms. Equal masses of the white solid are recovered when 30.0 mL of 0.10 M Na_2SO_4 solution is added to 70.0 mL of 0.20 M $Sr(NO_3)_2$ solution and when 30.0 mL of 0.20 M Na_2SO4 solution is added to 70.0 ml of 0.20 M $Sr(NO_3)_2$ solution.		

This Performance Expectation and associated Evidence Statements support the following Task Demands.	
Task Demands	
1. Make simple calculations using given data to estimate, calculate, and/or predict the masses of substances involved in a chemical reaction. These calculations may include the optimal ratio of reactants for a chemical reaction, mass of the limiting reactant, the mass of the excess reactant, theoretical yield, and actual yield.	
2. Illustrate, graph, describe, and/or identify the proportional relationships between substances involved in a chemical reaction that can be used to calculate or estimate the masses of atoms in the reactants and the products of that chemical reaction.	
3. Describe and predict simple chemical reactions in terms of mass, using proportional relationships among the substances involved in a chemical reaction.	
4. Compile, from given information, the particular data needed for a particular inference about the amounts of matter within a chemical system. This can include sorting out the relevant data from the given information.	

<table>
<tr><td>Performance Expectation</td><td colspan="3">HS-PS1-8
Develop models to illustrate the changes in the composition of the nucleus of the atom and the energy released during the processes of fission, fusion, and radioactive decay.</td></tr>
<tr><td>Dimensions</td><td>Developing and Using Models
• Develop a model based on evidence to illustrate the relationships between systems or between components of a system.</td><td>PS1.C: Nuclear Processes
• Nuclear processes, including fusion, fission, and radioactive decays of unstable nuclei, involve release or absorption of energy. The total number of neutrons plus protons does not change in any nuclear process.</td><td>Energy and Matter
• In nuclear processes, atoms are not conserved, but the total number of protons plus neutrons is conserved.</td></tr>
<tr><td>Clarifications and Content Limits</td><td colspan="3">Clarification Statements
• Emphasis is on simple qualitative models, such as pictures or diagrams, and on the scale of energy released in nuclear processes relative to other kinds of transformations.

Content Limits
• Assessment does not include quantitative calculation of energy released. Assessment is limited to alpha, beta, and gamma radioactive decays.</td></tr>
<tr><td>Science Vocabulary Students are Expected to Know</td><td colspan="3">Absorption, transformation, nuclear reaction, nucleus, decay rate, fission, fusion, neutron, nuclear mass, unstable, half-life, radioactive, radiation, alpha particle, alpha decay, beta particle, beta emission, gamma radiation, atomic number, atomic mass, proton, radioactive decay</td></tr>
<tr><td>Science Vocabulary Students are Not Expected to Know</td><td colspan="3">Nucleon(s), radioisotopes, positron, positron emission, electron capture, radioactive series, nuclear disintegration series, magic numbers, nuclear transmutations, particle accelerators, transuranium elements, radiometric dating, becquerel (Bq) unit, curie (Ci) unit, Geiger counter, radiotracer, critical mass, supercritical mass, nuclear reactor, ionizing radiation, nonionizing radiation, target nucleus, bombarding particle, nuclear process, nuclear stability, particle emission, rate of nuclear decay, spontaneous nuclear reaction</td></tr>
<tr><td colspan="4">Phenomena</td></tr>
<tr><td>Context/ Phenomena</td><td colspan="3">Some example phenomena for HS-PS1-8:
• Rocks from the Tuna Creek area of the Grand Canyon were tested and found to contain less lead (Pb) and more uranium (U) than rocks from the Elves Chasm area of the Grand Canyon.
• A brand new nuclear fuel rod containing 3% U-235 was used in a nuclear reactor in New Jersey for 18 months. When it was taken out the reactor, it was found to contain 0.8% U-235, 5.2% fission products, and 1.2% plutonium.
• Scientists in Dubna, Russia, after using a heavy ion accelerator to smash berkelium and detected atoms of elements 115 and 113 along with alpha particles.</td></tr>
<tr><td colspan="4">This Performance Expectation and associated Evidence Statements support the following Task Demands.</td></tr>
<tr><td colspan="4">Task Demands</td></tr>
<tr><td colspan="4">1. Select or identify from a collection of components, including distractors, the components needed to model the changes in nuclear composition and energy released during fission, fusion, and/or radioactive decay.</td></tr>
<tr><td colspan="4">2. Identify missing components, relationships, or other limitations of the model.</td></tr>
</table>

3. Describe, select, or identify the relationships among components of the nucleus and/or nuclear processes that explains the release or absorption of energy and/or the conservation of protons and neutrons.
4. Assemble or complete, from a collection of potential model components, an illustration or flow chart that is capable of representing a release or absorption of energy from a nuclear process. This does not include labeling an existing diagram.
5. Manipulate the components of a model to demonstrate the changes, properties, processes, and/or events that act to result in the phenomenon.

<table>
<tr><td>Performance Expectation</td><td colspan="3">HS-PS2-1
Analyze data to support the claim that Newton’s second law of motion describes the mathematical relationship among the net force on a macroscopic object, its mass, and its acceleration.</td></tr>
<tr><td>Dimensions</td><td>Analyzing and Interpreting Data
• Analyze data using tools, technologies, and/or models (e.g., computational, mathematical) in order to make valid and reliable scientific claims or determine an optimal design solution.</td><td>PS2.A: Forces and Motion
• Newton’s second law accurately predicts changes in the motion of macroscopic objects.</td><td>Cause and Effect
• Empirical evidence is required to differentiate between cause and correlation and make claims about specific causes and effects.</td></tr>
<tr><td>Clarifications and Content Limits</td><td colspan="3">Clarification Statements
• Examples of data could include tables or graphs of position or velocity as a function of time for objects subject to a net unbalanced force, such as a falling object, an object rolling down a ramp, or a moving object being pulled by a constant force.

Content Limits
• Assessment is limited to one-dimensional motion and to macroscopic objects moving at non-relativistic speeds.
• Stating the law or naming the law is not part of this PE.</td></tr>
<tr><td>Science Vocabulary Students are Expected to Know</td><td colspan="3">Velocity, acceleration, net force, friction, air resistance, impulse, vectors, slope, y-intercept</td></tr>
<tr><td>Science Vocabulary Students are Not Expected to Know</td><td colspan="3">Jerk, terminal velocity</td></tr>
<tr><td colspan="4">Phenomena</td></tr>
<tr><td>Context/ Phenomena</td><td colspan="3">The phenomenon for these PEs are the given data. Phenomena should describe the data set(s) to be given in terms of patterns or relationships to be found in the data, and the columns and rows of a hypothetical table presenting the data, even if the presentation is not tabular. The description of the phenomenon should describe the presentation format of the data (e.g., maps, tables, graphs, etc).

Some example phenomena for HS-PS2-1:
• Force is removed from two vehicles’ accelerator pedals. The vehicles’ positions over time are given.
• A water tank railcar is pulled by a train engine at constant speed and develops a leak allowing water to escape. The position and velocities of the water tank and train over time are given.
• A heavy model rocket rises a shorter distance than a lighter model rocket using the same type of engine. The position of each rocket over time is given.
• A falling skydiver’s velocity increases for several minutes and then reaches a maximum speed. The velocity of the skydiver over time is given.</td></tr>
<tr><td colspan="4">This Performance Expectation and associated Evidence Statements support the following Task Demands.</td></tr>
</table>

Task Demands
1. Organize and/or arrange (e.g., using illustrations and/or labels), make calculations, or summarize data to highlight trends, patterns, or correlations.
2. Generate/construct graphs, tables, or assemblages of illustrations and/or labels of data that document patterns, trends or relationships in the motion of a macroscopic object. This may include sorting out distractors.
3. Construct, state, or select a claim or propose a design solution based on the relationships identified in the data.
4. Use relationships identified in the data to predict the motion of and changes in the motion of macroscopic objects.
5. Identify patterns or evidence in the data that supports inferences about the motion of and changes in the motion of macroscopic objects.

<table>
<tr><td>Performance Expectation</td><td colspan="3">HS-PS2-2
Use mathematical representations to support the claim that the total momentum of a system of objects is conserved when there is no net force on the system</td></tr>
<tr><td>Dimensions</td><td>Using Mathematics and Computational Thinking
• Use mathematical representations of phenomena to describe explanations</td><td>PS2.A: Forces and Motion
• Momentum is defined for a particular frame of reference; it is the mass times the velocity of the object.
• If a system interacts with objects outside itself, the total momentum of the system can change; however, any such change is balanced by changes in the momentum of objects outside the system.</td><td>Systems and System Models
• When investigating or describing a system, the boundaries and initial conditions of the system need to be defined.</td></tr>
<tr><td>Clarifications and Content Limits</td><td colspan="3">Clarification Statements
• Emphasis is on the quantitative conservation of momentum in interactions and the qualitative meaning of this principle
• Students should not be deriving formulas but can be using and manipulating them

Content Limits
• Assessment is limited to systems of no more than two macroscopic bodies moving in one dimension.
• <u>Students do not need to know:</u>
  o How to use a derivation to show that momentum is conserved only when there is no net force.
  o How to derive formulas regarding conservation of momentum.
  o How to resolve vectors and apply the understanding that momentum must be conserved in all directions.
  o Newton’s Laws by name</td></tr>
<tr><td>Science Vocabulary Students are Expected to Know</td><td colspan="3">Friction, transfer, deceleration, frame of reference, net force, acceleration, velocity, internal, external, conversion, closed system, Newton’s Second Law, collision, vector</td></tr>
<tr><td>Science Vocabulary Students are Not Expected to Know</td><td colspan="3">Elastic collision, inelastic collision, inertial frame of reference</td></tr>
<tr><td colspan="4">Phenomena</td></tr>
<tr><td>Context/ Phenomena</td><td colspan="3">Some example phenomena for HS-PS2-2:
• A pool player hits a cue ball towards a stationary 8-ball. The cue ball collides with the 8-ball, causing the 8-ball to move. The 8-ball slows down until it comes to a rest 5 seconds after the collision.
• Two pool balls collide with each other and two soccer balls collide with each other. After the collision, the soccer balls come to a stop quicker than the pool balls.
• A pool player hits a cue ball towards a stationary 8-ball. The cue ball collides with the 8-ball. The velocity of the 8-ball 1 second after the collision is greater than the velocity of the 8-ball 2 seconds after the collision.</td></tr>
</table>

	• Two hockey pucks collide during an ice hockey practice. A player realizes that the two pucks take a long time to come to rest on the ice. After practice, he makes two street hockey pucks collide on pavement. The pucks come to a stop more quickly than the ones on the ice did.
This Performance Expectation and associated Evidence Statements support the following Task Demands.	
Task Demands	
1. Make simple calculations using given data to calculate or estimate the total momentum in the system OR the momentum of individual objects within the system.	
2. Illustrate, graph, or identify relevant features or data that can be used to calculate or estimate the total momentum in the system OR the momentum of individual objects within the system.	
3. Calculate or estimate properties or relationships between momentum and other forces based on data from one or more sources.	
4. Identify data or compile from given information, the information needed to support inferences about net force and/or how momentum is conserved within a system. This can include sorting out the relevant data from the given information.	

<table>
<tr><td>Performance Expectation</td><td colspan="3">HS-PS2-3
Apply scientific and engineering ideas to design, evaluate, and refine a device that minimizes the force on a macroscopic object during a collision.</td></tr>
<tr><td>Dimensions</td><td>Constructing Explanations and Designing Solutions
• Apply scientific ideas to solve a design problem, taking into account possible unanticipated effects.</td><td>PS2.A: Forces and Motion

• If a system interacts with objects outside itself, the total momentum of the system can change; however, any such change is balanced by changes in the momentum of objects outside the system.

ETS1.A: Defining and Delimiting an Engineering Problem

• Criteria and constraints also include satisfying any requirements set by society, such as taking issues of risk mitigation into account, and they should be quantified to the extent possible and stated in such a way that one can tell if a given design meets them. (secondary)

ETS1.C: Optimizing the Design Solution

• Criteria may need to be broken down into simpler ones that can be approached systematically, and decisions about the priority of certain criteria over others (trade-offs) may be needed. (secondary)</td><td>Cause and Effect
• Systems can be designed to cause a desired effect.</td></tr>
<tr><td>Clarifications and Content Limits</td><td colspan="3">Clarification Statements
• Examples of evaluation and refinement could include determining the success of the device at protecting an object from damage and modifying the design to improve it.
• Examples of a device could include a football helmet or a parachute.

Content Limits
• Assessment is limited to qualitative evaluations and/or algebraic manipulations</td></tr>
<tr><td>Science Vocabulary Students are Expected to Know</td><td colspan="3">Exert, acceleration, deceleration, impact, inertia, Newton’s First law, Newton’s Second Law, Newton’s Third Law of Motion, impact, drag, velocity, qualitative, criteria, theoretical model, optimal, deformation, impulse, tradeoff</td></tr>
<tr><td>Science Vocabulary Students are Not Expected to Know</td><td colspan="3">Rationale, aesthetics, consideration, representation, aspect, specification, critical, compressibility</td></tr>
<tr><td colspan="4">Phenomena</td></tr>
<tr><td>Context/ Phenomena</td><td colspan="3">Engineering standards are built around meaningful design problems rather than phenomena. For this performance expectation, the design problem and solutions replace phenomena.

Some example design problems for HS-PS2-3:</td></tr>
</table>

	• Bikers need to be both protected and have total visibility when riding. Design a helmet that protects the biker from collisions while maintaining awareness for his surroundings. • Phone screens can be easily broken if dropped on the ground. Design a phone case that protects the phone from collisions while maintaining functionality. • Design a material that can be implemented on a pool table, athletic field turf (fake grass), or miniature golf green to prevent wear and tear on the playing surface. • Design an instrument case so that the instrument will still be in good condition even if the case is subject to being dropped or rolled around.

This Performance Expectation and associated Evidence Statements support the following Task Demands.
Task Demands
1. Identify or assemble from a collection, including distractors, the relevant aspects of the problem, that with the given design solutions, if implemented, will resolve/improve the device by minimizing impact force.
2. Using the given information, select or identify the criteria against which the device or solution should be judged.
3. Using given data, propose/illustrate/assemble a potential device (prototype) or solution in order to minimize impact forces.
4. Using given information, select or identify constraints that the device or solution must meet.
5. Using a simulator, test a proposed prototype and evaluate the outcomes, potentially including proposing and testing modifications to the prototype.

Performance Expectation	**HS-PS2-4** Use mathematical representations of Newton's Law of Gravitation and Coulomb's Law to describe and predict the gravitational and electrostatic forces between objects.		
Dimensions	**Using Mathematics and Computational Thinking** • Use mathematical representations of phenomena to describe explanations.	**PS2.B: Types of Interactions** • Newton's law of universal gravitation and Coulomb's law provide the mathematical models to describe and predict the effects of gravitational and electrostatic forces between distant objects. • Forces at a distance are explained by fields (gravitational, electric, and magnetic) permeating space that can transfer energy through space. Magnets or electric currents cause magnetic fields; electric charges or changing magnetic fields cause electric fields.	**Patterns** • Different patterns may be observed at each of the scales at which a system is studied and can provide evidence for causality in explanations of phenomena.
Clarifications and Content Limits	**Clarification Statements** • Emphasis is on both quantitative and conceptual descriptions of gravitational and electric fields. **Content Limits** • Assessment is limited to systems with two objects. • Mathematical models can involve a range of linear and nonlinear functions including trigonometric functions, exponentials and logarithms, and computational tools for statistical analysis to analyze, represent, and model data.		
Science Vocabulary Students are Expected to Know	attraction, charge, conductor, electric charge, induced electric current, electric field, electromotive force, electromagnetic field, electromagnet, frequency, induction, insulator, magnetic field, magnetic field lines, magnetic force, permanent magnet, polarity, repulsion, resistance, voltage, battery, direction, magnitude, ampere, charged particle, volts, right-hand rule, tesla, vectors,		
Science Vocabulary Students are Not Expected to Know	electric potential, electromotive force, permeating, quantum property, Laplace force, electrodynamics, magnetic dipole, electrostatic, general relativity, Ampere's Law, Coulomb force, Lorentz force		
Phenomena			
Context/ Phenomena	Some example phenomena for HS-PS2-4: • A person weighs 150.00 pounds at sea level. At the top of Mount Everest, the same person weighs 149.25 pounds. • When an uncharged sphere is brought near another stationary uncharged sphere, the stationary uncharged sphere does not move. When a charged sphere is brought near a stationary uncharged sphere, the stationary uncharged sphere always moves towards the charged sphere. • The constant of proportionality in Coulomb's Law is 10^{20} times greater than the constant of proportionality in Newton's Law of Gravitation. However, the force of gravity on objects on Earth is usually much greater than the force exerted by magnets.		
This Performance Expectation and associated Evidence Statements support the following Task Demands.			
Task Demands			

8. Make simple calculations using given data to calculate or estimate the gravitational or electrostatic forces between objects.
9. Illustrate, graph, or identify relevant features or data that can be used to calculate or estimate the gravitational or electrostatic forces between objects.
10. Calculate or estimate gravitational or electrostatic properties/relationships based on data from one or more sources.
11. Compile, from given information, the particular data needed for a particular inference about the gravitational or electrostatic forces between objects. This can include sorting out the relevant data from the given information.

<table>
<tr><td>Performance Expectation</td><td colspan="3">HS-PS2-5
Plan and conduct an investigation to provide evidence that an electric current can produce a magnetic field and that a changing magnetic field can produce an electric current.</td></tr>
<tr><td>Dimensions</td><td>Planning and Carrying Out Investigations
• Plan and conduct an investigation individually and collaboratively to produce data to serve as the basis for evidence, and in the design: decide on types, how much, and accuracy of data needed to produce reliable measurements and consider limitations on the precision of the data (e.g., number of trials, cost, risk, time), and refine the design accordingly.</td><td>PS2.B: Types of Interactions
• Forces at a distance are explained by fields (gravitational, electric, and magnetic) permeating space that can transfer energy through space. Magnets or electric currents cause magnetic fields; electric charges or changing magnetic fields cause electric fields.

PS3.A: Definitions of Energy
• “Electrical energy” may mean energy stored in a battery or energy transmitted by electric currents. (secondary)</td><td>Cause and Effect
• Empirical evidence is required to differentiate between cause and correlation and make claims about specific causes and effects.</td></tr>
<tr><td>Clarifications and Content Limits</td><td colspan="3">Content Limits
• Assessment is limited to designing and conducting investigations with provided materials and tools.
• Coulomb Law is provided in the stimulus if student is required to make calculations.</td></tr>
<tr><td>Science Vocabulary Students are Expected to Know</td><td colspan="3">conductor, electric charge, induced electric current, electromotive force, electromagnetic field, electromagnet, induction, insulator, magnetic field, magnetic field lines, permanent magnet, polarity, resistance, voltage, magnitude, ampere, charged particle, volts, right-hand rule, tesla, vectors,</td></tr>
<tr><td>Science Vocabulary Students are Not Expected to Know</td><td colspan="3">electric potential, electromotive force, permeating, quantum property, Laplace force, electrodynamics, magnetic dipole, electrostatic, general relativity, Ampere’s Law, Coulomb force, Lorentz force</td></tr>
<tr><td colspan="4">Phenomena</td></tr>
<tr><td>Context/ Phenomena</td><td colspan="3">Some example phenomena for HS-PS2-5:
• Paperclips on a table are picked up by a wire when both ends of the wire are attached to a battery.
• When an electric current flows through a coil near a strong magnet, the coil rotates.
• The light bulb in a closed circuit turns on when a magnet moves near the wire in the circuit.
• A wind turbine built with a neodymium magnet produces more electricity than a wind turbine built with a ferrite magnet.</td></tr>
<tr><td colspan="4">This Performance Expectation and associated Evidence Statements support the following Task Demands.</td></tr>
<tr><td colspan="4">Task Demands</td></tr>
<tr><td colspan="4">1. Identify from a list, including distractors, the materials/tools/steps needed for an investigation to provide evidence that an electric current produces a magnetic field or that a changing magnetic field produces an electric current.</td></tr>
</table>

2. Identify the outcome data that should be collected in an investigation to provide evidence that an electric current produces a magnetic field or that a changing magnetic field produces an electric current.
3. Evaluate the sufficiency and limitations of data collected to explain the phenomenon.
4. Make and/or record observations about the magnetic field created by an electric current or the electric current created by a changing magnetic field.
5. Analyze, manipulate, interpret and/or communicate the data from an investigation to provide evidence that an electric current produces a magnetic field or that a changing magnetic field produces an electric current.
6. Explain or describe the causal processes that lead to the observed data.
7. Select, describe, or illustrate a prediction made by applying the findings from an investigation about electric currents and magnetic fields.

<table>
<tr><td>Performance Expectation</td><td colspan="3">HS-PS2-6
Communicate scientific and technical information about why the molecular-level structure is important in the functioning of designed materials.</td></tr>
<tr><td>Dimensions</td><td>Obtaining, Evaluating, and Communicating Information
• Communicate scientific information (e.g. about the process of development and the design and performance of a proposed process or system) in multiple formats (including orally, graphically, textually, and mathematically).</td><td>PS1.A: Structure and Properties of Matter
• The structure and properties of matter at the bulk scale are determined by electrical forces within and between atoms. (secondary)

PS2.B: Types of Interactions
• Attraction and repulsion between electric charges at the atomic scale explain the structure, properties, ad transformations of matter, as well as the contact forces between material objects.</td><td>Structure and Function
• Investigating or designing new systems or structures requires a detailed examination of the properties of different materials, the structures of different components, and connections of components to reveal its function and/or solve a problem.</td></tr>
<tr><td>Clarifications and Content Limits</td><td colspan="3">Clarification Statements
• Emphasis is on the attractive and repulsive forces that determine the functioning of the material.
• Examples could include why electrically conductive materials are often made of metal, flexible but durable materials are made up of long chained molecules, and pharmaceuticals are designed to interact with specific receptors.
• Assessment is limited to provided molecular structures of specific designed materials.

Content Limits
• <u>Students do not need to know:</u> specific molecular structures; specific names of synthetic materials such as vinyl, nylon, etc.</td></tr>
<tr><td>Science Vocabulary Students are Expected to Know</td><td colspan="3">Macroscopic, microscopic, electrical conductivity, long chained molecules, contact force, electron sharing, electron transfer, polymers, network material, surface tension, synthetic polymer, monomer, reactivity, intermolecular forces, charge, conductor, electric charge, insulator, permanent magnet, polarity, resistance, charged particle, ionic bond, covalent bond, hydrogen bond, ductile, malleable, friction</td></tr>
<tr><td>Science Vocabulary Students are Not Expected to Know</td><td colspan="3">electric potential, electromotive force, permeating, quantum property, Laplace force, electrodynamics, magnetic dipole, electrostatic, general relativity, Ampere’s Law, Coulomb force, Lorentz force, Van der Waals forces, organic molecules</td></tr>
<tr><td colspan="4" align="center">Phenomena</td></tr>
<tr><td>Context/ Phenomena</td><td colspan="3">Some example phenomena for HS-PS2-6:
• Zinc oxide was dissolved in water and the resulting solution was very difficult to stir. Upon the addition of a clear, amber colored liquid, the solution became much thinner and easier to stir.
• Water was spilled on two shirts. One shirt absorbed the water very quickly, leaving a large wet spot. On the other shirt, the water formed tiny spheres and bounced off, leaving the shirt dry.</td></tr>
</table>

	• A sample of cotton fabric was dyed with two different kinds of dye and then was washed several times to determine how well the dye stayed in the fabric. One dye faded over time, the other did not. • Food cooked in a bronze-colored pot cooked quickly and evenly. Food cooked in a silver-colored pot took longer and was not evenly cooked.

This Performance Expectation and associated Evidence Statements support the following Task Demands.
Task Demands
1. Analyze and interpret scientific evidence from multiple scientific/technical sources including text, diagrams, charts, symbols, mathematical representations that provide evidence that electrostatic forces on the atomic and molecular scale result in contact forces (e.g., friction, normal forces, stickiness) on the macroscopic scale.*
2. Identify relationships or patterns in scientific evidence to describe how electrostatic forces are related to properties of designed materials.
3. Identify and communicate evidence for how the structure and properties of matter and the types of interactions of matter at the atomic scale determine its function.
4. Synthesize an explanation for the function and properties of designed materials that incorporates the scientific evidence from multiple sources.*
5. Evaluate the validity, credibility, accuracy, relevancy and/or possible bias of scientific/technical sources.

<table>
<tr><td>Performance Expectation</td><td colspan="3">HS-PS3-1
Create a computational model to calculate the change in energy of one component in a system when the change in energy of the other component(s) and energy flows in and out of the system are known.</td></tr>
<tr><td>Dimensions</td><td>Using Mathematics and Computational Thinking
• Create a computational model or simulation of a phenomenon, designed device, process or system</td><td>PS3.A: Definitions of Energy
• Energy is a quantitative property of a system that depends on the motion and interactions of matter and radiation within that system. That there is a single quantity called energy is due to the fact that a system’s total energy is conserved, even as, within the system, energy is continually transferred from one object to another and between its various possible forms.

PS3.B: Conservation of Energy and Energy Transfer
• Conservation of energy means that the total change of energy in any system is always equal to the total energy transferred into or out of the system.
• Energy cannot be created or destroyed, but it can be transported from one place to another and transferred between systems.
• Mathematical expressions, which quantify how the stored energy in a system depends on its configuration (e.g., relative positions of charged particles, compression of a spring) and how kinetic energy depends on mass and speed, allow the concept of conservation of energy to be used to predict and describe system behavior.
• The availability of energy limits what can occur in any system.</td><td>Systems and System Models
• Models can be used to predict the behavior of a system, but these predictions have limited precision and reliability due to the assumptions and approximations inherent in models.</td></tr>
<tr><td>Clarifications and Content Limits</td><td colspan="3">Clarification Statements
• Emphasis is on explaining the meaning of mathematical expressions used in the model.

Content Limits
• Assessment is limited to:
o Basic algebraic expressions or computations
o Systems of two or three components
o Thermal energy, kinetic energy, and/or the energies in gravitational, magnetic and electric fields.
• <u>Students do not need to know:</u> detailed understanding of circuits or thermodynamics</td></tr>
<tr><td>Science Vocabulary Students are Expected to Know</td><td colspan="3">Mechanical energy, potential energy, conversion, kinetic energy, conduction, electrical circuit, electrical current, heat radiation, insulate, resistor, Volt, Amp, Ohm’s Law</td></tr>
<tr><td>Science Vocabulary Students are Not Expected to Know</td><td colspan="3">Entropy, second law of thermodynamics, thermodynamics, Stirling cycle, Carnot cycle, capacitor, inductance, inductor, Faradays law</td></tr>
</table>

Phenomena	
Context/ Phenomena	Some example phenomena for HS-PS3-1: • A block is attached to a spring and laid down on a table. The spring is stretched by pulling the block a certain distance. The spring is then released. As the block oscillates back and forth, the amplitude of each successive oscillation gets smaller until the block stops moving. • A light bulb is hooked up to an energy source. When a resistor is added in series to the circuit, the brightness of the light bulb dims. • Two metal pots are placed on a stove top. Pot 1 has a metal handle while Pot 2 has a rubber handle. The stove is turned on and the pots heat up. After 10 minutes, the handle on Pot 1 is much hotter than the handle on Pot 2. • A toy truck is placed at the top of a frictionless ramp. When it travels down the ramp it collides with a stationary toy truck sitting on a horizontal surface (with friction) at the bottom of the ramp. The truck at the bottom of the ramp then begins to move.

This Performance Expectation and associated Evidence Statements support the following Task Demands.
Task Demands
7. Make simple calculations using given data to calculate or estimate the amount of energy in certain components of the system.
8. Illustrate, graph, or identify relevant features or data that can be used to calculate or estimate how energy changes in one component of the system affect the energy changes in another component of the system OR how the flow of energy into and out of the system affects the energy change of each component within the system.
9. Calculate or estimate properties for, or the relationships between, each component of the system based on data from one or more sources.
10. Compile, from given information, the particular data needed for a particular inference about how energy changes in one component of the system affects the energy changes in another component of the system. This can include sorting out the relevant data from the given information.

<table>
<tr><td>Performance Expectation</td><td colspan="3">HS-PS3-2
Develop and use models to illustrate that energy at the macroscopic scale can be accounted for as a combination of energy associated with the motions of particles (objects) and energy associated with the relative positions of particles (objects).</td></tr>
<tr><td>Dimensions</td><td>Developing and Using Models
• Develop and use a model based on evidence to illustrate the relationships between systems or between components of a system</td><td>PS3.A: Definitions of Energy
• Energy is a quantitative property of a system that depends on the motion and interactions of matter and radiation within that system. That there is a single quantity called energy is due to the fact that a system's total energy is conserved, even as, within the system, energy is continually transferred from one object to another and between its various possible forms.
• At the macroscopic scale, energy manifests itself in multiple ways, such as in motion, sound, light and thermal energy.
• These relationships are better understood at the microscopic scale, at which all of the different manifestations of energy can be modeled as a combination of energy associated with the motion of particles and energy associated with the configuration (relative position of the particles). In some cases the relative position energy can be thought of as stored in fields (which mediate interactions between particles). This last concept includes radiation, a phenomenon in which energy stored in fields movers across space.</td><td>Energy and Matter
• Energy cannot be created or destroyed; it only moves between one place and another place, between objects and/or fields, or between systems.</td></tr>
<tr><td>Clarifications and Content Limits</td><td colspan="3">Clarification Statements
• Examples of phenomena at the macroscopic scale could include:
  • The conversion of kinetic energy to thermal energy
  • The energy stored due to position of an object above the Earth
  • The energy stored between two electrically-charged plates.
• Examples of models could include diagrams, drawings, descriptions, and computer simulations

Content Limits
• <u>Students do not need to know:</u>
  o Thermodynamics in detail
  o Gravitational fields
  o Thermonuclear fusion</td></tr>
<tr><td>Science Vocabulary Students are Expected to Know</td><td colspan="3">Mechanical energy, potential energy, kinetic energy, electric field, magnetic field, molecular energy, heat conduction, circuit, current, heat radiation, work</td></tr>
<tr><td>Science Vocabulary Students are</td><td colspan="3">Entropy, Second Law of Thermodynamics, thermodynamics, root mean velocity, Boltzmann's constant, gravitational fields, fusion, fission</td></tr>
</table>

Not Expected to Know	
Phenomena	
Context/ Phenomena	Some example phenomena for HS-PS3-2: • Two electrically charged plates, one with a positive charge and one with a negative charge, are placed a certain distance apart. Electron 1 is placed in the middle of the two plates. It accelerates to the positive plate and hits it with a certain velocity. Electron 2 is then placed closer to the negative plate. This electron gains more speed before reaching the positive plate. • A gas is placed inside a container and sealed with a piston. The outside of the container is heated up. The piston begins to move upwards. • A person rubs their hands together. Afterwards their hands feel warm. • A block is attached to a spring and placed on a horizontal table. When the spring is un-stretched, the spring and block do not move. When the spring is stretched to a certain distance (x), the block oscillates back and forth.
This Performance Expectation and associated Evidence Statements support the following Task Demands.	
Task Demands	
1. Select or identify from a collection of potential model components, including distractors, the components needed to model the phenomenon. Components might include equations used to calculate energy or objects used to set up the experimental model. The model can be a conceptual model (flow chart).	
2. Manipulate the components of a model to demonstrate how energy at the macroscopic scale can be accounted for as a combination of energy associated with the workings of particles at the microscopic scale, result in the observation of the phenomenon.	
3. Make predictions about the effects of changes in the motion or relative position of objects in the model. Predictions can be made by manipulating model components, completing illustrations, or selecting from lists with distractors.	
4. Identify missing components, relationships, or other limitations of the model showing how energy at the macroscopic scale is affected by the motion and positioning of particles at the microscopic scale.	
5. Describe, select, or identify the relationships among components of a model that describes, or explains, how energy is related to the motion and relative position of objects.	

<table>
<tr><td>Performance Expectation</td><td colspan="3">HS-PS3-3
Design, build and refine a device that works within given constraints to convert one form of energy into another form of energy.</td></tr>
<tr><td>Dimensions</td><td>Constructing Explanations and Designing Solutions
• Design, evaluate, and/or refine a solution to a complex real-world problem based on scientific knowledge, student-generated sources of evidence, prioritized criteria and tradeoff considerations.</td><td>PS3.A: Definitions of Energy
• At the macroscopic scale, energy manifests itself in multiple ways such as in motion, sound, light and thermal energy.

PS3.D: Energy in Chemical Processes
• Although energy cannot be destroyed, it can be converted to less useful forms – For example, to thermal energy in the surrounding environment.

ETS1.A: Defining and Delimiting an Engineering Problem
• Criteria and constraints also include satisfying any requirements set by society, such as taking issues of risk mitigation into account, and they should be quantified to the extent possible and stated in such a way that one can tell if a given design meets them. (secondary)</td><td>Energy and Matter
• Changes of energy and matter in a system can be described in terms of energy and matter flows into, out of, and within that system</td></tr>
<tr><td>Clarifications and Content Limits</td><td colspan="3">Clarification Statements
• Emphasis is on both qualitative and quantitative evaluations of devices.
• Examples of constraints could include use of renewable energy forms and efficiency.
• Examples of devices could include, but are not limited to:
    • Rube Goldberg devices
    • Wind Turbines
    • Solar cells
    • Solar ovens
    • Generators

Content Limits
• Assessment for quantitative evaluations is limited to total output for a given input. Assessment is limited to devices constructed with materials provided to students.</td></tr>
<tr><td>Science Vocabulary Students are Expected to Know</td><td colspan="3">Electric current, electrical energy, electromagnet, magnetic field, electric field, mechanical energy, renewable energy, generator, wind turbine, Rube Goldberg Device, solar cell, solar oven</td></tr>
<tr><td>Science Vocabulary Students are Not Expected to Know</td><td colspan="3">Torque, entropy, molecular energy, second law of thermodynamics, thermodynamics, thermal equilibrium, Stirling engine</td></tr>
<tr><td colspan="4">Phenomena</td></tr>
<tr><td>Context/ Phenomena</td><td colspan="3">Engineering standards are built around meaningful design problems rather than phenomena. For this performance expectation, the design problem and solutions replace phenomena.

Some example design problems for HS-PS3-3:
• Use and engine to generate the most light from an LED.</td></tr>
</table>

	• Refine a Stirling Engine to make it run for over 30mins. • Create a solar oven that will cook an egg in 10mins. • Refine a solar cell such that it maximizes energy output.

This Performance Expectation and associated Evidence Statements support the following Task Demands.
Task Demands
1. Articulate, describe, illustrate, or select the relationships, interactions, and/or processes to be explained. This may entail sorting relevant from irrelevant information or features.
2. Express or complete a causal chain explaining how the device converts one form of energy into another form of energy. This may include indicating directions of causality in an incomplete model such as a flow chart or diagram, or completing cause and effect chains.
3. Using given information, select or identify constraints that the energy converting device or solution must meet.
4. Identify evidence supporting the inference of causation that is expressed in a causal chain.
5. Using given data, propose, illustrate, or assemble a potential energy converting device (prototype) or solution.
6. Using a simulator, test a proposed energy converting prototype and evaluate the outcomes, potentially including proposing and testing modifications to the prototype.

Performance Expectation	**HS-PS3-4** Plan and conduct an investigation to provide evidence that the transfer of thermal energy when two components of different temperature are combined within a closed system results in a more uniform energy distribution among the components in the system (second law of thermodynamics).		
Dimensions	**Planning and Carrying Out Investigations** • Plan and conduct an investigation individually and collaboratively to produce data to serve as the basis for evidence, and in the design: decide on types, how much, and accuracy of data needed to produce reliable measurements and consider limitations on the precision of the data (e.g., number of trials, cost, risk, time), and refine the design accordingly.	**PS3.B: Conservation of Energy and Energy Transfer** • Energy cannot be created or destroyed, but it can be transported from one place to another and transferred between systems. • Uncontrolled systems always evolve toward more stable states—that is, toward more uniform energy distribution (e.g., water flows downhill, objects hotter than their surrounding environment cool down). **PS3.D: Energy in Chemical Processes** • Although energy cannot be destroyed, it can be converted to less useful forms—for example, to thermal energy in the surrounding environment.	**Systems and System Models** • When investigating or describing a system, the boundaries and initial conditions of the system need to be defined and their inputs and outputs analyzed and described using models.
Clarifications and Content Limits	**Clarification Statements** • Emphasis is on analyzing data from student investigations and using mathematical thinking to describe the energy changes both quantitatively and conceptually. • Examples of investigations could include mixing liquids at different initial temperatures or adding objects at different temperatures to water. **Content Limits** • *Assessment is limited to investigations based on materials and tools provided to students.*		
Science Vocabulary Students are Expected to Know	Specific heat, specific heat capacity, kinetic energy, microscopic scale, macroscopic scale, molecular energy, heat conduction, heat radiation, Kelvin, Joules, calorimetry		
Science Vocabulary Students are Not Expected to Know	Entropy, root mean velocity, Boltzmann’s constant, gravitational fields, fusion, fission		
Phenomena			
Context/ Phenomena	Some example phenomena for HS-PS3-4: • The temperature of a can of soda decreases when the can is placed in a container of ice. • Hot coffee cools down after cold cream is added to the cup. • A scoop of ice cream begins to melt when added to cold soda in a glass. • A foam cup has 200 grams of room temperature water after 100 grams of hot water are mixed with 100 grams of cold water.		

This Performance Expectation and associated Evidence Statements support the following Task Demands.	
Task Demands	
6. Identify, make, plan, and/or record observations/outcome data concerning changes in substances' properties in order to provide evidence of transfer of thermal energy within a closed system.	
7. Organize, arrange, and/or generate/construct graphs, flowcharts, tables, or assemblages of illustrations and/or labels of data that document patterns, trends, or correlations among observations and data concerning transfer of thermal energy within a closed system, and/or the boundaries of a closed system in which thermal energy is transferred.	
8. Describe, analyze, and/or summarize data (e.g., using illustrations and/or labels), to identify/highlight trends, perform calculations and other mathematical analyses, and identify patterns or correlations among observations and data concerning the transfer of thermal energy within a closed system.	
9. Use evidence to identify the boundaries of a closed system in which thermal energy is transferred.	
10. Identify patterns or evidence in the data that support inferences related to the transfer of thermal energy within a closed system.	

<table>
<tr><td>Performance Expectation</td><td colspan="3">HS-PS3-5
Develop and use a model of two objects interacting through electric or magnetic fields to illustrate the forces between objects and the changes in energy of the objects due to the interaction.</td></tr>
<tr><td>Dimensions</td><td>Developing and Using Models
• Develop and use a model based on evidence to illustrate the relationships between systems or between components of a system.</td><td>PS3.C: Relationship Between Energy and Forces
• When two objects interacting through a field change relative position, the energy stored in the field is changed.</td><td>Cause and Effect
• Cause and effect relationships can be suggested and predicted for complex natural and human-designed systems by examining what is known about smaller scale mechanisms within the system.</td></tr>
<tr><td>Clarifications and Content Limits</td><td colspan="3">Clarification Statements
• Examples of models could include: Drawings, diagrams, and texts, such as drawings of what happens when two charges of opposite polarity are near each other

Content Limits
• Assessment is limited to systems containing two objects
• <u>Students do not need to know:</u> Gauss’ Law, Ampere’s Law, Faraday’s Law or anything that requires in depth knowledge of the electromagnetism as a unified force.</td></tr>
<tr><td>Science Vocabulary Students are Expected to Know</td><td colspan="3">Electric current, acceleration, net force, newton’s second law of motion, inertia, velocity, magnet, electrical energy, magnetic force, attraction, repulsion, electromagnet, Coulomb’s law, electric/magnetic field, potential energy, kinetic energy</td></tr>
<tr><td>Science Vocabulary Students are Not Expected to Know</td><td colspan="3">Semiconductor, superconductor, torque, Gauss’ Law, Ampere’s Law, Lorentz force, Faraday’s Law, Lenz’s Law</td></tr>
<tr><td colspan="4">Phenomena</td></tr>
<tr><td>Context/ Phenomena</td><td colspan="3">Some example phenomena for HS-PS3-5:
• Two magnets are held close together such that they attract each other. When the magnets are further away from each other it is easier to keep them apart.
• A light bulb connected to a circuit with a battery lights up. When a stronger battery is placed in the circuit, the light bulb becomes brighter.
• A magnet rotates when placed in a magnetic field perpendicular to the magnet. When the magnet is brought close to the source of the magnetic field, it rotates faster.
• A water molecule is placed in an electric field. After it is released, it begins to rotate. After it rotates 90 degrees, it stops rotating.</td></tr>
<tr><td colspan="4">This Performance Expectation and associated Evidence Statements support the following Task Demands.</td></tr>
<tr><td colspan="4">Task Demands</td></tr>
<tr><td colspan="4">1. Select or identify from a collection of potential model components, including distractors, the components needed to model the phenomenon.</td></tr>
<tr><td colspan="4">2. Assemble or complete, from a collection of potential model components, an illustration or flow chart that is capable of representing how the forces between the objects and the energy of each object changes. This <u>does not</u> include labeling an existing diagram.*</td></tr>
</table>

3. Manipulate the components of a model to demonstrate the changes, properties, processes, and/or events that act to result in the phenomenon.*
4. Make predictions about the effects of changes in orientation of objects, distance between objects or size of magnetic and electric charges on the forces between objects and the energy of each object. Predictions can be made by manipulating model components, completing illustrations, or selecting from lists with distractors.**
5. Describe, select, or identify the relationships among components of a model that describe or explains the behavior of electric and magnetic fields and/or how that affects the forces between objects and the energy of the objects.
6. Identify missing components, relationships, or other limitations of the model.

*denotes those task demands which are deemed appropriate for use in stand-alone development
**TD 4 can only be used with TD2

<table>
<tr><td>Performance Expectation</td><td colspan="3">HS-PS4-1
Use mathematical representations to support a claim regarding relationships among the frequency, wavelength, and speed of waves traveling in various media.</td></tr>
<tr><td>Dimensions</td><td>Using Mathematics and Computational Thinking
• Use mathematical representations of phenomena or design solutions to describe and/or support claims and/or explanations.</td><td>PS4.A: Wave properties
• The wavelength and frequency of a wave are related to each other by the speed of travel of the wave, which depends on the type of wave and the media through which it is passing.</td><td>Cause and Effect
• Empirical evidence is required to differentiate between cause and correlation and make claims about specific causes and effects.</td></tr>
<tr><td>Clarifications and Content Limits</td><td colspan="3">Clarification Statements
• Examples of data could include electromagnetic radiation traveling in a vacuum and glass, sound waves traveling through air and water, and seismic waves traveling through Earth.

Content Limits
• Assessment is limited to algebraic relationships and describing those relationships qualitatively.
• Students are not expected to produce equations from memory, like Snell’s Law, but the concepts and relationships should be assessed.</td></tr>
<tr><td>Science Vocabulary Students are Expected to Know</td><td colspan="3">Simple wave, vacuum, electromagnetic radiation, radiation, wave source, index of refraction, Snell’s Law, angle of incidence, angle of reflection, normal at the point of incidence, critical angle, interface.</td></tr>
<tr><td>Science Vocabulary Students are Not Expected to Know</td><td colspan="3">Clausius–Mossotti relation, dielectric constant, Fermat’s principle, phase velocity, permittivity, permeability.</td></tr>
<tr><td colspan="4">Phenomena</td></tr>
<tr><td>Context/ Phenomena</td><td colspan="3">Some example phenomena for HS-PS4-1:
• A person uses their car horn in an effort to attract the attention of their friend who is swimming in a pool a short distance away. The friend hears only muffled noises.
• A person opens their curtains so that the sun shines in the window. A diamond in their necklace begins to sparkle brightly.
• An earthquake occurs in Japan. The vibrations are recorded in Brazil, but not in Miami.
• A person sees a fish through the glass wall of a rectangular fish tank. The person moves and looks through the end of the tank. The fish appears to be in a different place.</td></tr>
<tr><td colspan="4">This Performance Expectation and associated Evidence Statements support the following Task Demands.</td></tr>
<tr><td colspan="4">Task Demands</td></tr>
<tr><td colspan="4">1. Make calculations using given data to calculate or estimate relationships among the frequency, wavelength, speed of waves, and the media that they travel in.</td></tr>
<tr><td colspan="4">2. Illustrate, graph, or identify relevant features or data that can be used to calculate or estimate relationships among the frequency, wavelength, speed of waves, and the media that they travel in.</td></tr>
</table>

3. Calculate or estimate properties or relationships among the frequency, wavelength, and speed of waves in various media based on data from one or more sources.
4. Compile, from given information, the particular data needed for a particular inference about a relationship among the frequency, wavelength, speed of waves, and the media they travel in. This can include sorting out the relevant data from the given information.
5. Use quantitative or abstract reasoning to support a claim/explanation about a particular relationship between the velocity, wavelength, and frequency.

<table>
<tr><td>Performance Expectation</td><td colspan="3">HS-PS4-2
Evaluate questions about the advantages of using a digital transmission and storage of information.</td></tr>
<tr><td>Dimensions</td><td>Asking Questions and Defining Problems
• Evaluate questions that challenge the premise(s) of an argument, the interpretation of a data set, or the suitability of a design.</td><td>PS4.A: Wave Properties
• Information can be digitized (e.g., a picture stored as the values of an array of pixels); in this form, it can be stored reliably in computer memory and sent over long distances as a series of wave pulses.</td><td>Stability and Change
• Systems can be designed for greater or lesser stability.</td></tr>
<tr><td>Clarifications and Content Limits</td><td colspan="3">Clarification Statements
• Examples of advantages could include that digital information is stable because it can be stored reliably in computer memory, transferred easily, and copied and shared rapidly.
• Disadvantages could include issues of easy deletion, security, and theft.

Content Limits
• Assessment does not include the specific mechanism of any given device.</td></tr>
<tr><td>Science Vocabulary Students are Expected to Know</td><td colspan="3">Wave pulse, Wi-Fi device, binary, capacity, civilization, interdependence, degradation, emit, pixel, suitability, performance, analog, digital, progress, vacuum, electromagnetic radiation, computer, machine, radio wave, USB, bit, byte, discrete vs. continuous, decode, encode.</td></tr>
<tr><td>Science Vocabulary Students are Not Expected to Know</td><td colspan="3">Analog jack, HDMI, router, impedance, granularity, bandwidth.</td></tr>
<tr><td colspan="4">Phenomena</td></tr>
<tr><td>Context/ Phenomena</td><td colspan="3">Some example phenomena for HS-PS4-2:
• A person uses e-mail to back up all of their personal data.
• A person is reading some science papers that were written in 1905 and wonders how people got so much great research done before the internet was invented.
• One day in June 2009 a person noticed that their old analog television stopped broadcasting their favorite television channel.
• A person stays in constant contact with all of their friends and relatives using their cell phone.</td></tr>
<tr><td colspan="4">This Performance Expectation and associated Evidence Statements support the following Task Demands.</td></tr>
<tr><td colspan="4">Task Demands</td></tr>
<tr><td colspan="4">1. Identify or construct an empirically testable question(s) based on advantages and disadvantages associated with the phenomenon. In addition to other plausible distractors, distractors may include non-testable (“nonscientific”) questions.</td></tr>
<tr><td colspan="4">2. Make and/or record observations about the factors that affect digitally stored or transmitted data.</td></tr>
<tr><td colspan="4">3. Assemble or complete an illustration, flow chart, or graph based on an empirically testable question that is capable of identifying clear advantages or disadvantages associated with digital transmission and storage of information in the phenomenon.</td></tr>
</table>

4. Select or describe conclusions relevant to a question posed and supported by the data, especially inferences about causes and effects.
5. Make predictions about the phenomenon derived from the questions. Predictions can be made by manipulating model components, completing illustrations, or selecting from lists with distractors. Predict outcomes when properties are changed, given the inferred cause and effect relationships.
6. Compile, from given information, the particular data needed for a particular inference about the advantages/disadvantages. This can include sorting out the relevant data from the given information.

<table>
<tr><td>Performance Expectation</td><td colspan="3">HS-PS4-3
Evaluate the claims, evidence and reasoning behind the idea that electromagnetic radiation can be described either by a wave model or a particle model, and that for some situations one model is more useful than the other.</td></tr>
<tr><td>Dimensions</td><td>Engaging in Argument from Evidence
• Evaluate the claims, evidence and reasoning behind currently accepted explanations or solutions to determine the merits of arguments</td><td>PS4.A: Wave Properties
• Waves can add or cancel one another as they cross, depending on their relative phase (i.e. relative position of peaks and troughs of waves), but they emerge unaffected by each other. (Boundary: the discussion at this grade level is qualitative only; it can be based on the fact that two different sounds can pass a location in different directions without getting mixed up).

PS4.B: Electromagnetic Radiation
• Electromagnetic radiation (e.g., radio, microwaves, light) can be modeled as a wave of changing electric and magnetic fields or as particles called photons. The wave model is useful for explaining many features of electromagnetic radiation, and the particle model explains other features.</td><td>Systems and System Models
• Models (e.g., physical, mathematical and computer models) can be used to simulate systems and interactions – including energy, matter and information flows – within and between systems at different scales.</td></tr>
<tr><td>Clarifications and Content Limits</td><td colspan="3">Clarification Statements
• Emphasis is on how experimental evidence supports the claim and how a theory is generally modified in light of new evidence. Assessment should only test the qualitative aspect of the wave model vs. particle model.
• Examples of a phenomenon could include:
  • Resonance
  • Interference
  • Diffraction
  • Photoelectric Effect

Content Limits
• Assessment does not include using Quantum Theory
• Assessment does not include in depth calculations
• <u>Students do not need to know:</u> Specific types of electromagnetic radiation and their wavelengths/frequencies</td></tr>
<tr><td>Science Vocabulary Students are Expected to Know</td><td colspan="3">Interference, diffraction, refraction, photoelectric effect, emission, absorption, brightness, resonance, transmission, visible light, transverse wave</td></tr>
<tr><td>Science Vocabulary Students are Not Expected to Know</td><td colspan="3">Doppler effect for light (redshift), microwave radiation, ultraviolet radiation, ionize, infrared radiation, wave-particle duality, quantum, quanta, x-ray, gamma rays, radio waves, oscillations, electrostatic induction, Planck’s equation, Planck’s constant, magnetic dipole, electric dipole,</td></tr>
<tr><td colspan="4">Phenomena</td></tr>
<tr><td>Context/ Phenomena</td><td colspan="3">Some example phenomena for HS-PS4-3:</td></tr>
</table>

	• When light hits a metal, a stream electrons are ejected from the metal. When the color of the light pointed at the metal changes, the kinetic energy of the stream of electrons changes. • Light is made to pass through two small slits on a piece of dark construction paper. The light that goes through the slits is then projected on a second piece of dark of construction. A pattern of bright and dark bands is seen on the second piece of dark construction paper. • The emission spectra of Hydrogen is completely black but for 4 discrete lines violet, blue, green and red color. • A red laser is pointed at a glass prism. The light bends as it goes through the prism. A violet laser is then pointed at the glass prism and the light bends more than the light from the red laser.

This Performance Expectation and associated Evidence Statements support the following Task Demands.
Task Demands
1. Based on the provided data or information, identify the explanation that describes light behaves like a particle and or behaves like a wave.
2. Identify and/or explain the claims, evidence, and reasoning supporting the explanation that light can behave like a particle or a wave, and why certain evidence is best explained by only one of these models.
3. Identify and/or describe additional relevant evidence not provided that would support or clarify the explanation of how light can behave like a particle or a wave.
4. Evaluate the strengths and weaknesses of a claim to explain which pieces of evidence support the fact that light behaves as a particle or a wave.
5. Analyze and/or interpret evidence and its ability to support the explanation that light can behave as both a wave and a particle.
6. Provide and/or evaluate reasoning to support the explanation that light can behave as both a wave and a particle and that some evidence is only supported by one of the models.

<table>
<tr><td>Performance Expectation</td><td colspan="3">HS-PS4-4
Evaluate the validity and reliability of claims in published materials of the effects that different frequencies of electromagnetic radiation have when absorbed by matter.</td></tr>
<tr><td>Dimensions</td><td>Obtaining, Evaluating, and Communicating Information
• Evaluate the validity and reliability of multiple claims that appear in scientific and technical texts or media reports, verifying the data when possible.</td><td>PS4.B: Electromagnetic Radiation
• Electromagnetic radiation (e.g., radio, microwaves, light) can be modeled as a wave of changing electric and magnetic fields or as particles called photons. The wave model is useful for explaining many features of electromagnetic radiation, and the particle model explains other features.</td><td>Cause and Effect
• Cause and effect relationships can be suggested and predicted for complex natural and human designed systems by examining what is known about smaller scale mechanisms within the system.</td></tr>
<tr><td>Clarifications and Content Limits</td><td colspan="3">Clarification Statements
• Emphasis is on the idea that photons associated with different frequencies of light have different energies, and the damage to living tissue from electromagnetic radiation depends on the energy of the radiation.
• Examples of published materials could include trade books, magazines, web resources, videos, and other passages that may reflect bias.

Content Limits
• Assessment is limited to qualitative descriptions.</td></tr>
<tr><td>Science Vocabulary Students are Expected to Know</td><td colspan="3">Interference, diffraction, refraction, photoelectric effect, emission, absorption, brightness, resonance, photon, visible light, transverse wave, phase, transparent, light scattering, light transmission, radio wave, visible light, electric potential, gamma ray, infrared radiation, ionize, microwave, ohm, photoelectric, ultraviolet,</td></tr>
<tr><td>Science Vocabulary Students are Not Expected to Know</td><td colspan="3">Doppler effect for light (redshift), microwave radiation, ultraviolet radiation, infrared radiation, wave-particle duality, quantum, quanta, x-ray, gamma rays, oscillations, electrostatic induction, Planck’s equation, Planck’s constant, magnetic dipole, electric dipole</td></tr>
<tr><td colspan="4">Phenomena</td></tr>
<tr><td>Context/ Phenomena</td><td colspan="3">Some example phenomena for HS-PS4-4:
• A student places a glass bowl filled with soup in a microwave. After a minute in the microwave, the soup is hotter than the glass bowl.
• A lit candle is placed at one end of a tube filled with carbon dioxide. A student standing at the other end of the tube can see the candle’s flame. When looking through a monitor that looks at the infrared radiation emitted by the flame, the student can no longer see the candle’s flame.
• Astronauts aboard the International Space Station are exposed to a different amount of ultraviolet radiation from the sun than humans on Earth.
• In 2020, NASA is sending a rover to Mars with multiple materials on it in order to test whether or not they can be used as space suits for future Mars travelers. Orthofabric was chosen to be sent on the mission, while Spectra was not.</td></tr>
<tr><td colspan="4">This Performance Expectation and associated Evidence Statements support the following Task Demands.</td></tr>
</table>

Task Demands
1. Analyze and/or interpret scientific evidence from multiple scientific/technical sources including text, diagrams, charts, symbols, mathematical representations that provide evidence of the effects that different frequencies of electromagnetic radiation have when absorbed by matter.*
2. Identify relationships or patterns in scientific evidence to describe how different frequencies of electromagnetic radiation effect matter when absorbed.
3. Illustrate, graph, or identify relevant features or data that can be used to communicate information about the effect that different frequencies of electromagnetic radiation have on matter when it is absorbed.
4. Synthesize an explanation for the effects of electromagnetic radiation on matter when absorbed that incorporates the scientific evidence from multiple sources.*
5. Evaluate the validity, credibility, accuracy, relevancy and/or possible bias of scientific/technical sources.
6. Identify the cause and effect reasoning in a claim from the sources, including the extrapolations to larger scales from cause and effect relationships of mechanisms at small scales (e.g. extrapolating from the effect of a particular wavelength of radiation on a single cell to the effect of that wavelength on the entire organism).

*denotes those task demands which are deemed appropriate for use in stand-alone item development

Performance Expectation	**HS-PS4-5** Communicate technical information how some technological devices use the principles of wave behavior and wave interaction with matter to transmit and capture information and energy.		
Dimensions	**Obtaining, Evaluating, and Communicating Information** • Communicate technical information or ideas (e.g., about phenomena and/or the process of development and the design and performance of a proposed process or system) in multiple formats (including orally, graphically, textually, and mathematically).	**PS3.D: Energy in Chemical Processes** • Solar cells are human-made devices that likewise capture the sun's energy and produce electrical energy. *(secondary)* **PS4.A: Wave Properties** • Information can be digitalized (e.g., a picture stored as the values of an array of pixels); in this form, it can be stored reliably in computer memory and sent over long distances as a series of wave pulses. **PS4.B: Electromagnetic Radiation** • Photoelectric materials emit electrons when they absorb light of a high enough frequency. **PS4.C: Information Technologies** • Multiple technologies based on the understanding of waves and their interactions with matter are part of everyday experiences in the modern world (e.g., medical imaging, communications, scanners) and in scientific research. They are essential tools for producing, transmitting, and capturing signals and for storing and interpreting the information contained in them.	**Cause and Effect** • Systems can be designed to cause a desired effect.
Clarifications and Content Limits	**Clarification Statements** • Examples could include solar cells capturing light and converting it to electricity, medical imaging, and communications technology. **Content Limits** • Assessments are limited to qualitative information. • Assessment does not include band theory.		
Science Vocabulary Students are Expected to Know	refraction, reflection, infrared, electromagnetic spectrum, constructive wave, destructive wave, restoring, periodic motion, mechanical wave, interference, velocity, diffraction, standing wave, nodes, angle of incidence, rarefaction, superposition, medium, longitudinal wave, transverse wave, standing wave, ultrasound, dispersion, intensity, prism, resonance, radar, sonar, virtual image, real image		
Science Vocabulary Students are Not Expected to Know	Constructive interference, destructive interference, light ray, total internal reflection		
Phenomena			
Context/ Phenomena	Engineering Standards are built around meaningful design problems rather than phenomena. For this performance expectation, the design problem and solutions replace phenomena.		

	Some example design problems and/or solutions for HS-PS4-5: • When using light detection and ranging (LiDAR) over a forested area the light reflects off multiple surfaces and affects the accuracy of elevation models. • Solar cells only capture about 20% of the energy from the sun. • Marine radar is mounted to the front of ships used for collision avoidance. Occasionally the radar can distort the coast line and report a straight coastline when it is curved. • Water reflects radar, blanking out entire regions of radar screens.

This Performance Expectation and associated Evidence Statements support the following Task Demands.
Task Demands
1. Analyze and interpret scientific evidence from multiple scientific/technical sources including text, diagrams, charts, symbols, mathematical representations that provide evidence of how devices use wave behavior, the absorption of photons, and the production of electrons to solve problems.
2. Identify relationships or patterns in scientific evidence to describe how waves are used to produce, transmit, and capture signals in electronic devices.
3. Illustrate, graph, or identify relevant features or data that can be used to communicate wave information and
4. Synthesize an explanation for the function and properties of designed materials that incorporates the scientific evidence from multiple sources.
5. Evaluate the validity, credibility, accuracy, relevancy and/or possible bias of scientific/technical sources.

<table>
<tr><td>Performance Expectation</td><td colspan="3">HS-LS1-1
Construct an explanation based on evidence for how the structure of DNA determines the structure of proteins, which carry out the essential functions of life through systems of specialized cells.</td></tr>
<tr><td>Dimensions</td><td>Constructing Explanations and Designing Solutions
• Construct an explanation based on valid and reliable evidence obtained from a variety of sources (including students’ own investigations, models, theories, simulations, peer review) and the assumption that theories and laws that describe the natural world operate today as they did in the past and will continue to do so in the future.</td><td>LS1.A: Structure and Function
• Systems of specialized cells within organisms help them perform the essential functions of life.
• All cells contain genetic information in the form of DNA molecules. Genes are regions in the DNA that contain the instructions that code for the formation of proteins, which carry out most of the work of cells.</td><td>Structure and Function
• Investigating or designing new systems or structures requires a detailed examination of the properties of different materials, the structures of different components, and the connections of these components in order to solve problems.</td></tr>
<tr><td>Clarifications and Content Limits</td><td colspan="3">Content Limits
• Assessment does not include identification of specific cell or tissue types, whole body systems, specific protein structures and functions, or the biochemistry of protein synthesis.</td></tr>
<tr><td>Science Vocabulary Students are Expected to Know</td><td colspan="3">Nucleus, chromosome, DNA, nucleated cell, transcription, double helix, adenine, guanine, cytosine, thymine, deoxyribose, phosphate, hydrogen bond, nucleotide base, mRNA, amino acid, translation</td></tr>
<tr><td>Science Vocabulary Students are Not Expected to Know</td><td colspan="3">primary, secondary, tertiary protein structure, tRNA, ribosome.</td></tr>
<tr><td colspan="4">Phenomena</td></tr>
<tr><td>Context/ Phenomena</td><td colspan="3">Sample phenomena for HS-LS1-1:
• Sweat glands cool the body by releasing sweat onto the skin’s surface. A protein transports salt to help carry the water to the skin’s surface. In some individuals, the salt is not reabsorbed and is left on the skin.
• When a blood vessel is cut, several proteins act to form a blood clot. This blood clot helps to stop the loss of blood from the body. In some individuals, when a blood vessel is cut, a blood clot does not form.
• During cell division, a copy of DNA in the cell is made. Sometimes mistakes are made in the DNA copy that are corrected by specific proteins. In some cells, when those mistakes in the DNA are not corrected, uncontrolled cellular division results.
• After a person eats, sugars from food are absorbed from the bloodstream into the body’s cells. Insulin—a polypeptide hormone—allows those cells to absorb glucose from the bloodstream. In some individuals, sugars are not absorbed into the body’s cells and are left in the bloodstream.</td></tr>
<tr><td colspan="4">This Performance Expectation and associated Evidence Statements support the following Task Demands.</td></tr>
<tr><td colspan="4">Task Demands</td></tr>
</table>

9. Describe the cause and effect relationship between a DNA sequence and the structure/function of a protein. This may include indicating the directions of causality in a model or completing a cause and effect chain.
10. Describe, identify, or select evidence that supports or contradicts a claim about the role of DNA in causing the phenomenon. The evidence may be obtained from valid source(s) or might be generated by students using a simulation.
11. Given an appropriate explanation for a phenomenon, predict the effects of subsequent changes to a DNA sequence in protein structure and function. Predictions may be selected from a collection of possibilities, including distractors, or they might be illustrated or described in writing.
12. Use evidence to construct an explanation of how protein structure and subsequent function depend on a DNA sequence.
13. Identify and justify additional pieces of evidence that would help distinguish between competing hypotheses.

<table>
<tr><td>Performance Expectation</td><td colspan="3">HS-LS1-2
Develop and use a model to illustrate the hierarchical organization of interacting systems that provide specific functions within multicellular organisms.</td></tr>
<tr><td>Dimensions</td><td>Developing and Using Models
• Develop and use a model based on evidence to illustrate the relationships between systems or between components of a system.</td><td>LS1.A: Structure and Function
▪ Multicellular organisms have a hierarchical structural organization, in which any one system is made up of numerous parts and is itself a component of the next level.</td><td>Systems and System Models
• Models (e.g., physical, mathematical, and computer models) can be used to simulate systems and interactions — including energy, matter, and information flows — within and between systems at different scales.</td></tr>
<tr><td>Clarifications and Content Limits</td><td colspan="3">Clarification Statements
• Emphasis is on functions at the organism system level such as nutrient uptake, water delivery, and organism movement in response to neural stimuli. An example of an interacting system could be an artery depending on the proper function of elastic tissue and smooth muscle to regulate and deliver the proper amount of blood within the circulatory system.

Content Limits
• Assessment does not include interactions and functions at the molecular or chemical reaction level (e.g., hydrolysis, oxidation, reduction, etc.).
• Assessment does not include mutations in genes that could contribute to modified bodily functions.</td></tr>
<tr><td>Science Vocabulary Students Are Expected to Know</td><td colspan="3">Circulatory, respiratory, digestive, excretory, nervous, immune, integumentary, skeletal, muscle, reproductive, external stimuli, cell, tissue, organ,</td></tr>
<tr><td>Science Vocabulary Students Are Not Expected to Know</td><td colspan="3">Synaptic transmission, neuron, neurotransmitter, biofeedback, hormonal signaling.</td></tr>
<tr><td colspan="4">Phenomena</td></tr>
<tr><td>Context/ Phenomena</td><td colspan="3">Some example phenomena for HS-LS1-2:
• After a healthy person eats a large meal, both their blood pressure and heart rate increase.
• When a normal adult male exercises, both his breathing rate and heart rate increase.
• The area around a person’s skin where a small scab has formed feels warm to the touch.
• Skin surface capillaries dilate when a person feels hot.</td></tr>
<tr><td colspan="4">This Performance Expectation and associated Evidence Statements support the following Task Demands.</td></tr>
<tr><td colspan="4">Task Demands</td></tr>
<tr><td colspan="4">1. Assemble or complete an illustration or flow chart that is capable of representing how structures in two (or more) body systems interact to carry out normal, necessary bodily functions. This <u>does not</u> include labeling an existing diagram.*</td></tr>
</table>

2. Using the developed model, identify and describe the relationships between the structures and their coordinated functions in two (or more) body systems.
3. Using the developed model, show that interacting systems have a hierarchical organization and provide specific functions within the body at those specific levels or organization.*
4. Make predictions about, or generate explanations for, how additions/substitutions/removal of certain components in the model can interrupt or change the relationships between those components and, thus, the bodily functions carried out by those structures in two (or more) body systems.
5. Given models or diagrams of hierarchical organization of interacting systems, identify the components and the mechanism in each level of the hierarchy OR identify the properties of the components that allow those functions to occur.
6. Identify missing components, relationships, or other limitations of the model.

*denotes those task demands which are deemed appropriate for use in stand-alone item development

<table>
<tr><td>Performance Expectation</td><td colspan="3">HS-LS1-3
Plan and conduct an investigation to provide evidence that feedback mechanisms maintain homeostasis.</td></tr>
<tr><td>Dimensions</td><td>Planning and Carrying Out Investigations
• Plan and conduct an investigation individually and collaboratively to produce data to serve as the basis for evidence. In the design decide on types, how much, and accuracy of data needed to produce reliable measurements and consider limitations on the precision of the data (e.g., number of trials, cost, risk, time), and refine the design accordingly.</td><td>LS1.A: Structure and Function
• Feedback mechanisms maintain a living system’s internal conditions within certain limits and mediate behaviors, allowing it to remain alive and functional even as external conditions change within some range. Feedback mechanisms can encourage (through positive feedback) or discourage (negative feedback) what is going on inside the living system.</td><td>Stability and Change
• Feedback (negative or positive) can stabilize or destabilize a system.</td></tr>
<tr><td>Clarifications and Content Limits</td><td colspan="3">Clarification Statements
• Examples of investigations could include heart rate response to exercise, stomate response to moisture and temperature, and root development in response to water levels.

Content Limits
• Assessment does not include the cellular processes involved in the feedback mechanism.</td></tr>
<tr><td>Science Vocabulary Students Are Expected to Know</td><td colspan="3">Equilibrium, steady state, stable state, balanced state, stimulus, receptor, biotic factor, abiotic factor, external environment, internal environment, muscle, nerve, hormone, enzyme, chemical regulator, gland, system, metabolism, disturbance, fluctuation, maintenance, concentration, hibernation, convection, conduction, radiation, evaporation.</td></tr>
<tr><td>Science Vocabulary Students Are Not Expected to Know</td><td colspan="3">Effector, osmoregulation, conformer, set point, sensor, circadian rhythm, acclimatization, thermoregulation, endothermic, ectothermic, integumentary system, countercurrent exchange, bioenergetics, basal metabolic rate, standard metabolic rate, torpor, poikilotherm, homeotherm, countercurrent heat exchange.</td></tr>
<tr><td colspan="4">Phenomena</td></tr>
<tr><td>Context/ Phenomena</td><td colspan="3">Some example phenomena for HS-LS1-3:
• Fruit ripeness (positive feedback loop):
  o In nature, a tree or bush will suddenly ripen all of its fruits or vegetables without any visible signal.
• Human blood sugar concentration (negative feedback loop):
  o The liver both stores and produces sugar in response to blood glucose concentration.
  o The pancreas releases either glucagon or insulin in response to blood glucose concentration.
• Sunning lizards (negative feedback loop):
  o Lizards sun on a warm rock to regulate body temperature.
• Thermoregulation in dolphins due to counter-current arrangement of veins around arteries (negative feedback loop):
  o The counter-current system minimizes the loss of heat incurred when blood travels to the different parts of dolphins’ bodies.</td></tr>
</table>

This Performance Expectation and associated Evidence Statements support the following Task Demands.
Task Demands
1. Identify the outcome data that should be collected in an investigation to provide evidence that feedback mechanisms maintain homeostasis. This could include measurements and/or identifications of changes in the external environment, the response of the living system, stabilization/destabilization of the system's internal conditions, and/or the number of systems for which data are collected.
2. Make and/or record observations about the external factors affecting systems interacting to maintain homeostasis, responses of living systems to external conditions, and/or stabilization/destabilization of the systems' internal conditions.*
3. Identify or describe the relationships, interactions, and/or processes that contribute to and/or participate in the feedback mechanisms maintaining homeostasis that lead to the observed data.
4. Using the collected data, express or complete a causal chain explaining how the components of (a) mechanism(s) interact in response to a disturbance in equilibrium in order to maintain homeostasis. This may include indicating directions of causality in an incomplete model such as a flow chart or diagram, or completing cause and effect chains.*
5. Evaluate the sufficiency and limitations of data collected to explain the cause and effect mechanism(s) maintaining homeostasis.

*denotes those task demands which are deemed appropriate for use in stand-alone item development

<table>
<tr><td>Performance Expectation</td><td colspan="3">HS-LS1-4
Use a model to illustrate the role of cellular division (mitosis) and differentiation in producing and maintaining complex organisms.</td></tr>
<tr><td>Dimensions</td><td>Developing and Using Models
• Use a model based on evidence to illustrate the relationships between systems or between components of a system.</td><td>LS1.B: Growth and Development of Organisms
• In multicellular organisms, individual cells grow and then divide via a process called mitosis, thereby allowing the organism to grow. The organism begins as a single cell (fertilized egg) that divides successively to produce many cells, with each parent cell passing identical genetic material (two variants of each chromosome pair) to both daughter cells. Cellular division and differentiation produce and maintain a complex organism, composed of systems of tissues and organs that work together to meet the needs of the whole organism.</td><td>Systems and System Models
• Models (e.g., physical, mathematical, computer models) can be used to simulate systems and interactions—including energy, matter, and information flows—within and between systems at different scales.</td></tr>
<tr><td>Clarifications and Content Limits</td><td colspan="3">Clarification Statements
• Assessment does not include specific gene control mechanisms or rote memorization of the steps of mitosis.

Content Limits
• <u>Students do not need to know</u>: Specific names of the stages of mitosis – Interphase, G1 phase, S phase, G2 phase, prophase, metaphase, anaphase, telophase, cytokinesis.</td></tr>
<tr><td>Science Vocabulary Students Are Expected to Know</td><td colspan="3">Nucleus, chromosome, sister chromatids, sperm cell, egg cell, fertilize, genome, gene, differential gene expression, cellular differentiation, cellular division, cytoplasm, daughter cell, parent cell, somatic cell, cell cycle, homologous, haploid, diploid, DNA.</td></tr>
<tr><td>Science Vocabulary Students Are Not Expected to Know</td><td colspan="3">Spindle, metaphase plate, cleavage furrow, chromatin modification, transcription regulation initiation, enhancers, transcription factors, post-transcriptional regulation; noncoding RNAs, cytoplasmic determinants, inductive signals, chiasmata, kinetochore, microtubule.</td></tr>
<tr><td colspan="4" align="center">Phenomena</td></tr>
<tr><td>Context/ Phenomena</td><td colspan="3">Some example phenomena for HS-LS1-4:
• Genomic sequencing of a parent cell and one of its daughter cells reveals that both have the same genetic makeup.
• At the end of an hour, approximately 30,000 skin cells were shed by a person, but a loss of mass was not noticeable.
• Ears and noses can be grown from stem cells in laboratory.
• Plant cells in a root tip longitudinal cross section are different sizes and shapes.</td></tr>
<tr><td colspan="4" align="center">This Performance Expectation and associated Evidence Statements support the following Task Demands.</td></tr>
<tr><td colspan="4" align="center">Task Demands</td></tr>
<tr><td colspan="4">1. Assemble or complete an illustration or flow chart that is capable of representing how a parent (somatic) cell is formed through fertilization, undergoes cellular division, forming daughter cells, and how those daughter cells</td></tr>
</table>

	contain all genetic material from the parent cells but differentiate via gene expression necessary. This does **not** include labeling an existing diagram.*
2.	Using the model, identify and describe the relationship between the amount and composition of the genetic material that daughter cells receive from parent cells.
3.	Using the model, show that in multicellular organisms, different cell types arise from differential gene expression, not because of dissimilar genetic material within the cell's nucleus.
4.	Use a model of cellular division and differentiation to explain/illustrates the relationships between components that allow multicellular organisms to grow and carry out specific and necessary functions.*
5.	Given models or diagrams of cellular division and differentiation, show that cells form tissues and organs that have specific structures and interact to carry out specific and necessary functions.
6.	Identify missing components, relationships, or other limitations of the model.

*denotes those task demands which are deemed appropriate for use in stand-alone item development

<table>
<tr><td>Performance Expectation</td><td colspan="3">HS-LS1-5
Use a model to illustrate how photosynthesis transforms light energy into stored chemical energy.</td></tr>
<tr><td>Dimensions</td><td>Developing and Using Models
• Use a model based on evidence to illustrate the relationship between systems or between components of a system.</td><td>LS1.C: Organization for Matter and Energy Flow in Organisms
• The process of photosynthesis converts light energy to stored chemical energy by converting carbon dioxide plus water into sugars plus released oxygen.</td><td>Energy and Matter
• Changes of energy and matter in a system can be described in terms of energy and matter flows into, out of, and within that system.</td></tr>
<tr><td>Clarifications and Content Limits</td><td colspan="3">Clarification Statements
• Emphasis is on illustrating inputs and outputs of matter and the transfer and transformation of energy in photosynthesis by plants and other photosynthesizing organisms.
• Examples of models could include diagrams, chemical equations, and conceptual models.

Content Limits
• Assessment does not include specific biochemical steps or cell signaling pathways.</td></tr>
<tr><td>Science Vocabulary Students are Expected to Know</td><td colspan="3">Organic, hydrocarbon, net transfer, chloroplast, chlorophyll, cytoplasm, mitochondria, vacuole, nucleus, protein, ATP, amino acid, autotroph(s), heterotroph(s), algae</td></tr>
<tr><td>Science Vocabulary Students are Not Expected to Know</td><td colspan="3">Thylakoid, NADP(H/$^+$), Calvin cycle, carbon fixation, redox reactions, electron transport chain, oxidative phosphorylation, photoautotroph(s), mesophyll, stomata, stroma, thylakoids, thylakoid membrane, light reactions, carotenoids, cytochrome complex, C_3 pants, C_4 plants</td></tr>
<tr><td colspan="4">Phenomena</td></tr>
<tr><td>Context/ Phenomena</td><td colspan="3">Some example phenomena for HS-LS1-5:
• A maple tree in Washington state survives in the winter after losing all of its leaves.
• The waters of the Laguna Grande lagoon in Puerto Rico give off a bluish-green glow at night when disturbed.
• On the sill of a stained glass window, a soy plant behind the red glass panel grew taller than a soy plant behind the green glass panel.
• In a parking lot in the city of Bordeaux, France a tank filled with algae produces a green light.</td></tr>
<tr><td colspan="4">This Performance Expectation and associated Evidence Statements support the following Task Demands.</td></tr>
<tr><td colspan="4">Task Demands</td></tr>
<tr><td colspan="4">1. Assemble or complete, from a collection of potential model components and distractors, an illustration or flow chart that is capable of representing the transformation of light energy into stored chemical energy.</td></tr>
<tr><td colspan="4">2. Use a model to identify and describe the relationships in terms of matter and/or energy between the reactants and the products of photosynthesis.*</td></tr>
</table>

3. Use a model to show the transfer of matter and flow of energy between an organism and its environment during photosynthesis.*
4. Make predictions about how additions/substitutions/removals of model components affect the transformation of light energy into stored chemical energy.*
5. Sort relevant from irrelevant information to support a model that demonstrates how sugar and oxygen are produced by carbon dioxide and water through the process of photosynthesis.
6. Given models or diagrams of photosynthesis, identify the components and the mechanism in each scenario OR identify the properties of the components that allow photosynthesis to occur.*
7. Identify missing components, relationships, or other limitations of a model intended to show how photosynthesis transforms light energy into stored chemical energy.
8. Describe changes of energy and matter that occur in a system due to photosynthesis.

*denotes those task demands which are deemed appropriate for use in stand-alone item development

<table>
<tr><td>Performance Expectation</td><td colspan="3">HS-LS1-6
Construct and revise an explanation based on evidence for how carbon, hydrogen, and oxygen from sugar molecules may combine with other elements to form amino acids and/or other large carbon-based molecules.</td></tr>
<tr><td>Dimensions</td><td>Constructing Explanations and Designing Solutions
• Construct and revise an explanation based on valid and reliable evidence obtained from a variety of sources (including students’ own investigations, models, theories, simulations, peer review) and the assumption that theories and laws that describe the natural world operate today as they did in the past and will continue to do so in the future.</td><td>LS1.C: Organization for Matter and Energy Flow in Organisms
• Sugar molecules formed contain carbon, hydrogen, and oxygen. Their hydrocarbon backbones are used to make amino acids and other carbon-based molecules that can be assembled into larger molecules (such as proteins or DNA), used, for example, to form new cells.
• As matter and energy flow through different organizational levels of living systems, chemical elements are recombined in different ways to form different products.</td><td>Energy and Matter
• Changes of energy and matter in a system can be described as energy and matter flowing into, out of, and within that system.</td></tr>
<tr><td>Clarifications and Content Limits</td><td colspan="3">Clarification Statements
• Emphasis is on using evidence from models and simulations to support explanations.

Content Limits
• Assessment does not include the details of the specific chemical reactions or identification of macromolecules.
• <u>Students do not need to know:</u> Specific biochemical pathways and processes. Specific enzymes, oxidation-reduction</td></tr>
<tr><td>Science Vocabulary Students Are Expected to Know</td><td colspan="3">Hydrocarbon, carbohydrate, amino acid, nucleic acid, DNA, ATP, lipid, fatty acid, ingestion, rearrangement, stable, open system.</td></tr>
<tr><td>Science Vocabulary Students Are Not Expected to Know</td><td colspan="3">Endothermic reaction, exothermic reaction, aerobic respiration, oxidation, reduction, oxidation-reduction reaction, glycolysis, citric acid cycle, electron transport chain.</td></tr>
<tr><td colspan="4" align="center">Phenomena</td></tr>
<tr><td>Context/ Phenomena</td><td colspan="3">Some example phenomena for HS-LS1-6:
• Hagfish produce and are covered in a thick layer of protective slime.
• The black widow spider’s silk is several times as strong as any other known spider silk, making it about as durable as Kevlar.
• The female silk moth, releases a pheromone that is sensed by the male’s feather-like antennae, inducing his excited fluttering behavior.
• The bombardier beetle release a boiling, noxious, pungent spray that can repel potential predators.</td></tr>
</table>

This Performance Expectation and associated Evidence Statements support the following Task Demands.
Task Demands
1. Describe, identify, or select evidence supporting or contradicting a claim that sugar molecules containing organic elements (e.g., carbon, hydrogen, and oxygen) that are ingested by an organism are broken down and rearranged via chemical reactions to form proteins, lipids, and nucleic acids.
2. Identify and justify additional pieces of evidence that would help distinguish between competing hypotheses.
3. Express or complete a description of the flow of energy and/or matter within and between living systems. This may include indicating directions of causality in an incomplete model such as a flow chart or diagram, or completing cause and effect chains.*
4. Articulate, describe, or select the relationships, interactions, reactions and/or processes to be explained. This may entail sorting relevant from irrelevant information or features of the reactants and products.*
5. Given an appropriate explanation for a phenomenon, predict the effects of subsequent changes in the amount and types of organic molecules ingested and the amount and type of products formed within a living system.

*denotes those task demands which are deemed appropriate for use in stand-alone item development

<table>
<tr><td>Performance Expectation</td><td colspan="3">HS-LS1-7
Use a model to illustrate that cellular respiration is a chemical process whereby the bonds of food molecules and oxygen molecules are broken and the bonds in new compounds are formed, resulting in a net transfer of energy.</td></tr>
<tr><td>Dimensions</td><td>Developing and Using Models
• Use a model based on evidence to illustrate the relationships between systems or between components of a system.</td><td>LS1.C: Organization for Matter and Energy Flow in Organisms
• As matter and energy flow through different organizational levels of living systems, chemical elements are recombined in different ways to form different products
• As a result of these chemical reactions, energy is transferred from one system of interacting molecules to another. Cellular respiration is a chemical process in which the bonds of food molecules and oxygen molecules are broken and new compounds are formed that can transport energy to muscles. Cellular respiration also releases the energy needed to maintain body temperature despite ongoing energy transfer to the surrounding environment.</td><td>Energy and Matter
• Energy cannot be created or destroyed—it only moves between one place and another, between objects and/or fields, or between systems.</td></tr>
<tr><td>Clarifications and Content Limits</td><td colspan="3">Clarification Statement
• Emphasis is on the conceptual understanding of the inputs and outputs of the process of cellular respiration.

Content Limits
• Students aren’t expected to identify the steps or specific processes involved in cellular respiration.
• Assessment does not include mechanisms of cellular respiration (enzymatic activity, oxidation, molecular gradients, etc.).
• <u>Students do not need to know:</u> enzymes, ATP synthase, metabolism, biochemical pathways, redox reactions, molecular transport.</td></tr>
<tr><td>Science Vocabulary Students Are Expected to Know</td><td colspan="3">ATP, chemical bonds, energy transfer, glycolysis, enzymes, mitochondria, cytosol, cytoplasm, nitrogen, adenine, phosphate, amino acid.</td></tr>
<tr><td>Science Vocabulary Students Are Not Expected to Know</td><td colspan="3">Biochemical, fatty acids, oxidizing agent, electron acceptor, biosynthesis, locomotion, phosphorylation, electron transport chain, chemiosmosis, pyruvate, pentose.</td></tr>
<tr><td colspan="4">Phenomena</td></tr>
<tr><td>Context/ Phenomena</td><td colspan="3">Some example phenomena for HS-LS1-7:
• A young plant is grown in a bowl of sugar water. As it grows, the amount of sugar in the water decreases.</td></tr>
</table>

	• A bacterial colony in a petri dish is continually provided with sugar water. Over the course of a few days, the bacteria grow larger. When sugar water is no longer provided, the colonies shrink and some disappear. • A person feels tired and weak before eating lunch. After eating some fruit, the person feel more energetic and awake. • An athlete completing difficult training feels that her muscles recover and repair faster when she eats more food in a day, compared to when she ate less food in a day.
This Performance Expectation and associated Evidence Statements support the following Task Demands.	
Task Demands	
14. Assemble or complete an illustration or flow chart that is capable of representing the transformation of food plus oxygen into energy and/or new compounds. This *does not* include labeling an existing diagram.	
15. Using the developed model, identify and describe the relationships between the reactants of the transformation and the products of the transformation.*	
16. Using the developed model, show that matter and energy are only rearranged during cellular respiration, but never created or destroyed.	
17. Make predictions about how additions/substitutions/removals of certain components can maintain/destroy the balance of the food plus oxygen → energy/new compounds reaction.*	
18. Given models or diagrams of cellular respiration, identify the components and the mechanism in each scenario OR identify the properties of the components that allow cellular respiration to occur.	
19. Identify missing components, relationships, or other limitations of the model.	
20. Describe, select, or identify the relationships among components of a model that describe or explain cellular respiration.	

*denotes those task demands which are deemed appropriate for use in stand-alone item development

<table>
<tr><td>Performance Expectation</td><td colspan="3">HS-LS2-1
Use mathematical and/or computational representations to support explanations of factors that affect carrying capacity of ecosystems at different scales.</td></tr>
<tr><td>Dimensions</td><td>Using Mathematical and Computational Thinking
• Use mathematical and/or computational representations of phenomena or design solutions to support explanations</td><td>LS2.A: Interdependent Relationships in Ecosystems
▪ Ecosystems have carrying capacities, which are limits to the numbers of organisms and populations they can support. These limits result from such factors as the availability of living and nonliving resources and from challenges such as predation, competition and disease. Organisms would have the capacity to produce populations of greater size were it not for the fact that environments and resources are finite. This fundamental tension affects the abundance (number of individuals) of species in any given ecosystem.</td><td>Scale, Proportion, and Quantity
• The significance of a phenomenon is dependent on the scale, proportion, and quantity involved.</td></tr>
<tr><td>Clarifications and Content Limits</td><td colspan="3">Clarification Statements
• Emphasis is on quantitative analysis and comparison of the relationships among interdependent factors, including boundaries, resources, climate, and competition.
• Examples of mathematical comparisons could include graphs, charts, histograms, and population changes gathered from simulations or historical data sets.
• Examples of mathematical representations include finding the average, determining trends, and using graphic comparisons of multiple sets of data.

Content Limits
• Assessment does not include deriving mathematical equations to make comparisons.
• <u>Students do not need to know:</u> Calculus/advanced mathematics (e.g., exponential growth and decay).</td></tr>
<tr><td>Science Vocabulary Students Are Expected to Know</td><td colspan="3">Predation, interdependent, disturbance, equilibrium of ecosystems, fluctuation, stable, biotic, abiotic, sustain, anthropogenic, overexploitation, urbanization, population, emigrants, immigrants, exponential, generation, rebounding, limiting resources, logistic, competition, negative feedback, population control.</td></tr>
<tr><td>Science Vocabulary Students Are Not Expected to Know</td><td colspan="3">Dispersion, demography, survivorship curve (J or S), reproductive table, semelparity, iteroparity, metapopulation, demographic transition, resource partitioning, Shannon diversity, biomanipulation, density dependent selection (K-selection), density independent selection (r selection), intrinsic factors.</td></tr>
<tr><td colspan="4" align="center">Phenomena</td></tr>
<tr><td>Context/ Phenomena</td><td colspan="3">Some example phenomena for HS-LS2-1:
• On Ngorogoro Crater in Tanzania in 1963, a scientist sees that there are much fewer lions than there were on previous visits.
• On St. Matthew Island, reindeer were introduced in 1944, but today no reindeer can be found on the island.
• In Washington State, more harbor seals are present today than in the past.
• In Alaska, you can see many more brown bears in Lake Clark National Park than in Denali National Park.</td></tr>
</table>

This Performance Expectation and associated Evidence Statements support the following Task Demands.	
Task Demands	
1. Make calculations using given data to calculate or estimate factors affecting the carrying capacity of an ecosystem.*	
2. Illustrate, graph, or identify relevant features or data that can be used to calculate or estimate factors affecting the carrying capacity of ecosystems of different scales.*	
3. Calculate or estimate properties of or relationships between factors affecting the carrying capacity of an ecosystem based on data from one or more sources.	
4. Compile, from given information, the data needed for a particular inference about factors affecting the carrying capacity of an ecosystem. This can include sorting out the relevant data from the given information and representing the data through graphs, charts, and/or histograms.	
5. Use quantitative or abstract reasoning to make a claim about the factors that affect the carrying capacity of an ecosystem.	

*denotes those task demands which are deemed appropriate for use in stand-alone item development

<table>
<tr><td>Performance Expectation</td><td colspan="3">HS-LS2-2
Use mathematical representations to support and revise explanations, based on evidence about factors affecting biodiversity and populations in ecosystems of different scales.</td></tr>
<tr><td>Dimensions</td><td>Using Mathematical and Computational Thinking
• Use mathematical representations of phenomena or design solutions to support and revise explanations.</td><td>LS2.A: Interdependent Relationships in Ecosystems
• Ecosystems have carrying capacities, which are limits to the numbers of organisms and populations they can support. These limits results from factors such as the availability of living and nonliving resources and from such challenges such as predation, competition, and disease. Organisms would have the capacity to produce populations of greater size were it not for the fact that environments and resources are finite. This fundamental tension affects the abundance (number of individuals) of species in any given ecosystem.

LS2.C: Ecosystem Dynamics, Functioning, and Resilience
• A complex set of interactions within an ecosystem can keep its numbers and types of organisms relatively constant over long periods of time under stable conditions. If a modest biological or physical disturbance to an ecosystem occurs, it may return to its more or less original status (i.e., the ecosystem is resilient) as opposed to becoming a very different ecosystem. Extreme fluctuations in conditions or the size of any population, however, can challenge the functioning of ecosystems in terms of resources and habitat availability.</td><td>Scale, Proportion, and Quantity
• Using the concept of orders of magnitude allows one to understand how a model at one scale relates to a model at another scale.</td></tr>
<tr><td>Clarifications and Content Limits</td><td colspan="3">Clarification Statements
• Examples of mathematical representations include finding the average, determining trends, and using graphic comparisons of multiple sets of data.

Content Limits
• Assessment is limited to provided data.
• <u>Students do not need to know</u>: Calculus/advanced mathematics (e.g., exponential growth and decay)</td></tr>
<tr><td>Science Vocabulary Students Are Expected to Know</td><td colspan="3">Carrying capacity, anthropogenic changes, overexploitation, extinction, demographic, population pyramid, deforestation, habitat fragmentation, sustainable, abiotic factor, biotic factor, species richness, symbiosis, niche, fragile ecosystem, biodiversity index, zero population growth, density, dispersion, immigration, emigration, limiting factor</td></tr>
<tr><td>Science Vocabulary</td><td colspan="3">Water regime, direct driver, eutrophication, species evenness, range of tolerance, realized niche, niche generalist, niche specialist, edge habitat, endemic species, logistic growth model, exponential</td></tr>
</table>

Students Are Not Expected to Know	population growth, mark-recapture method, territoriality, demography, cohort, survivorship curve, reproductive table, life history, semelparity, iteroparity, K-selection, r-selection, dieback.
Phenomena	
Context/ Phenomena	Some example phenomena for HS-LS2-2: • After brown tree snakes were accidentally introduced to Guam in the 1950s, 11 native bird species went extinct. • When European settlers decreased the wolf population for safety, deer populations thrived and overconsumed native plant species. • California's Central Valley can support fewer waterfowl in the winter during drought. • The cones of Lodgepole pines do not release their seeds until a fire melts the resin that keeps them sealed.
This Performance Expectation and associated Evidence Statements support the following Task Demands.	
Task Demands	
1. Make simple calculations using given data to calculate or estimate factors affecting biodiversity and populations in ecosystems.	
2. Illustrate, graph, or identify relevant features or data that can be used to calculate or estimate factors affecting biodiversity and populations in ecosystems of different scales.	
3. Calculate or estimate properties of or relationships between factors affecting biodiversity and populations in ecosystems based on data from one or more sources.	
4. Compile, from given information, the data needed for a particular inference about factors affecting biodiversity and populations in ecosystems. This can include sorting out the relevant data from given information.	
5. Construct an explanation regarding the relationship between biodiversity and populations in ecosystems of different scales using the given, calculated, or compiled information.	
6. Revise or evaluate a given explanation of the relationship between biodiversity and populations in ecosystems of different scales based on the given, calculated, or compiled information.	

Performance Expectation	**HS-LS2-3** Construct and revise an explanation based on evidence for the cycling of matter and flow of energy in aerobic and anaerobic conditions.		
Dimensions	**Constructing Explanations and Designing Solutions** • Construct and revise an explanation based on valid and reliable evidence obtained from a variety of sources (including students' own investigations, models, theories, simulations, and peer review) and the assumption that theories and laws that describe the natural world operate today as they did in the past and will continue to do so in the future.	**LS2.B: Cycles of Matter and Energy Transfer in Ecosystems** • Photosynthesis and cellular respiration (including anaerobic processes) provide most of the energy for the processes.	**Energy and Matter** • Energy drives the cycling of matter within and between systems.
Clarifications and Content Limits	**Clarification Statements** • Emphasis is on conceptual understanding of the role of aerobic and anaerobic respiration in different environments. • Emphasis is on conceptual understanding that the supply of energy and matter restricts a system's operation; for example, without inputs of energy (sunlight) and matter (carbon dioxide and water), a plant cannot grow. **Content Limits** • Assessment does not include the specific chemical processes of either aerobic or anaerobic respiration. • Students do not need to know: lactic acid vs. alcoholic fermentation, chemical equations for photosynthesis, cellular respiration, or fermentation.		
Science Vocabulary Students Are Expected to Know	Organic compound synthesis, net transfer, biomass, carbon cycle, solar energy		
Science Vocabulary Students Are Not Expected to Know	Lactic acid fermentation, alcoholic fermentation, glycolysis, Kreb's cycle, electron transport chain.		
Phenomena			
Context/ Phenomena	Some example phenomena for HS-LS2-3: • After running for a long period of time, human muscles develop soreness and a burning sensation, and breathing rate increases. • Bread baked with yeast looks and tastes differently than bread that is baked without yeast. • A plant that is watered too much will have soft, brown patches on their leaves and will fail to grow. • Cyanobacteria differ from other bacteria in that cyanobacteria appear blue-green in color and also lack flagella.		
This Performance Expectation and associated Evidence Statements support the following Task Demands.			
Task Demands			

1. Describe, identify, or select evidence supporting or contradicting a claim about the role of photosynthesis and aerobic and anaerobic respiration in the cycling of matter and energy in an ecosystem.
2. Identify and justify additional pieces of evidence that would help distinguish between competing hypotheses.
3. Express or complete a description of the flow of energy and/or matter between organisms. This may include indicating directions of causality in an incomplete model such as a flow chart or diagram, or completing cause-and-effect chains.*
4. Articulate, describe, or select the relationships, interactions, and/or processes to be explained. This may entail sorting relevant from irrelevant information or features of the reactants and products.*
5. Given an appropriate explanation for a phenomenon, predict the effects of subsequent changes in environmental conditions on the flow of matter and energy between organisms.

<table>
<tr><td>Performance Expectation</td><td colspan="3">HS-LS2-4
Use mathematical representations to support claims for the cycling of matter and flow of energy among organisms in an ecosystem.</td></tr>
<tr><td>Dimensions</td><td>Using Mathematical and Computational Thinking
• Use mathematical representations of phenomena, or design solutions to support claims.</td><td>LS2.B: Cycles of Matter and Energy Transfer in Ecosystems
▪ Plants or algae from the lowest level of the food web. At each link upward in a food web, only a small fraction of the matter consumed at the lower level is transferred upward, to produce growth and release energy in cellular respiration at the higher level. Given this inefficiency, there are generally fewer organisms at higher levels of a food web. Some matter reacts to release energy for life functions, some matter is stored in newly made structures and much is discarded. The chemical elements that make up the molecules of organisms pass through food webs and into and out of the atmosphere and soil, and they are combined and recombined in different ways. At each link in an ecosystem, matter and energy are conserved.</td><td>Energy and Matter
• Energy cannot be created or destroyed; it only moves between one place and another place, between objects and/or fields, or between systems.</td></tr>
<tr><td>Clarifications and Content Limits</td><td colspan="3">Clarification Statements
• Emphasis is on using a mathematical model of stored energy in biomass to describe the transfer of energy from one trophic level to another, and that matter and energy are conserved as matter cycles and energy flows through ecosystems.
• Emphasis is on atoms and molecules—such as carbon, oxygen, hydrogen, and nitrogen—being conserved as they move through an ecosystem.

Content Limits
• Assessment is limited to proportional reasoning to describe the cycling of matter and flow of energy.
• <u>Students do not need to know</u>: the specific biochemical mechanisms or thermodynamics of cellular respiration to produce ATP or of photosynthesis to convert sunlight energy into glucose.</td></tr>
<tr><td>Science Vocabulary Students Are Expected to Know</td><td colspan="3">Interdependent, nutrient, hydrocarbon, transfer system, equilibrium of ecosystems, decomposer, producer, ATP, solar energy, predator-prey relationship, trophic level</td></tr>
<tr><td>Science Vocabulary Students Are Not Expected to Know</td><td colspan="3">Detritivore, denitrification, thermodynamics, nitrogen fixation, biogeochemical cycle, anaerobic process.</td></tr>
<tr><td colspan="4">Phenomena</td></tr>
<tr><td>Context/ Phenomena</td><td colspan="3">Some example phenomena for HS-LS2-4:
• In the 6,000-hectare rainforest of San Lorenzo, Panama, there are 312 arthropods for every mammal, including humans.</td></tr>
</table>

	• In Silver Springs, Florida, the biomass of plants is 809 g/m^2, while the biomass of large fish is 5 g/m^2. • A herd of grazing caribou in the Seward Peninsula of Alaska are seen eating the leaves of birch trees in July. In December, they are seen eating tree lichen. • A pine tree growing in a forest remains in one location throughout its lifetime. A fox in the same forest moves around every day of its life.

This Performance Expectation and associated Evidence Statements support the following Task Demands.
Task Demands
1. Calculate or estimate changes or differences in matter and energy between trophic levels of an ecosystem. **
2. Illustrate, graph, or identify a mathematical model describing changes in stored energy through trophic levels of an ecosystem.**
3. Compile and interpret data from given information to establish the relationship between organisms at different trophic levels.*
4. Use quantitative or abstract reasoning to make a claim about the cycling of matter and flow of energy through the trophic levels of an ecosystem. This may include sorting relevant from irrelevant information.*
5. Identify and describe the components of a mathematical representation of an ecosystem that could include relative quantities related to organisms, matter, energy, and the food web of that ecosystem.

*denotes those task demands which are deemed appropriate for use in stand-alone item development
**TDs 1 and 2 may be used for stand-alones in combination with TD3 and TD4.

<table>
<tr><td>Performance Expectation</td><td colspan="3">HS-LS2-5
Develop a model to illustrate the role of photosynthesis and cellular respiration in the cycling of carbon among the biosphere, atmosphere, hydrosphere, and geosphere.</td></tr>
<tr><td>Dimensions</td><td>Developing and Using Models
• Develop a model based on evidence to illustrate the relationships between systems or components of a system.</td><td>LS2.B: Cycles of Matter and Energy Transfer in Ecosystems
▪ Photosynthesis and cellular respiration are important components of the carbon cycle, in which carbon is exchanged among the biosphere, atmosphere, oceans, and geosphere through chemical, physical, geological, and biological processes.

PS3.D: Energy in Chemical Processes
▪ The main way that solar energy is captured and stored on Earth is through the complex chemical process known as photosynthesis. (secondary)</td><td>Systems and System Models
• Models (e.g., physical, mathematical, or computer models) can be used to simulate systems and interactions—including energy, matter, and information flows—within and between systems at different scales.</td></tr>
<tr><td>Clarifications and Content Limits</td><td colspan="3">Clarification Statements
• Examples of models could include simulations and mathematical models.

Content Limits
• Assessment does not include the specific chemical steps of photosynthesis and respiration.</td></tr>
<tr><td>Science Vocabulary Students Are Expected to Know</td><td colspan="3">Recycle, consumer, transform, organism, convert, decomposer, producer, hydrocarbon, microbes, ATP</td></tr>
<tr><td>Science Vocabulary Students Are Not Expected to Know</td><td colspan="3">Endothermic reaction, exothermic reaction, free energy, hydrolysis, oxidation.</td></tr>
<tr><td colspan="4">Phenomena</td></tr>
<tr><td>Context/ Phenomena</td><td colspan="3">Some example phenomena for HS-LS2-5:
• A herd of cows grazing in a field wear balloon-like backpack devices on their backs.
• A piece of coal preserving a fossil leaf imprint is burned within the furnace of a coal-fired electrical power plant. Smoke generated from the fire escapes out of a smoke stack
• Several acres of trees are cut down and burned, generating clouds of smoke.
• Two mice die in the woods in November, one in Massachusetts and one in Florida. The Florida mouse decomposes much more quickly than the Massachusetts mouse.</td></tr>
<tr><td colspan="4">This Performance Expectation and associated Evidence Statements support the following Task Demands.</td></tr>
<tr><td colspan="4">Task Demands</td></tr>
<tr><td colspan="4">1. Assemble or complete an illustration or flow chart that is capable of representing how the processes of photosynthesis and cellular respiration cycle carbon by various chemical, physical, geological, and biological</td></tr>
</table>

	processes through two or more spheres (biosphere, atmosphere, hydrosphere, geosphere). This *does not* include labeling an existing diagram.
2.	Using the developed model, identify and describe the relationships between the processes of photosynthesis and cellular respiration, and the coordinated functions of transferring carbon among two or more spheres (biosphere, atmosphere, hydrosphere, geosphere).
3.	Using the developed model, show that photosynthesis and cellular respiration are important parts of the overall carbon cycle that transfers carbon through two or more spheres (biosphere, atmosphere, hydrosphere, geosphere).
4.	Make predictions about, or generate explanations for, how substitutions of certain components in the model can interrupt or change the relationships between, or functions of, those components, thus effecting the cycling of carbon through the various spheres (biosphere, atmosphere, hydrosphere, geosphere).
5.	Given models or diagrams* of the processes of photosynthesis and cellular respiration, identify the components and the mechanisms in each process that cycle carbon OR identify the properties of the components that allow those functions to occur.
6.	Identify missing components, relationships, or other limitations of the model.
7.	Modify/augment/add to the model to change or add steps that can alter the cycling of carbon through the various spheres (biosphere, atmosphere, hydrosphere, and/or geosphere).

*Labeled diagrams by themselves are not usually sufficient to serve as models.

<table>
<tr><td>Performance Expectation</td><td colspan="3">HS-LS2-6
Evaluate the claims, evidence, and reasoning that the complex interactions in ecosystems maintain relatively consistent numbers and types of organisms in stable conditions, but changing conditions may result in a new ecosystem.</td></tr>
<tr><td>Dimensions</td><td>Engaging in Argument from Evidence

• Evaluate the claims, evidence, and reasoning behind currently accepted explanations or solutions to determine the merits of arguments.</td><td>LS2.C: Ecosystem Dynamics, Functioning, and Resilience

• A complex set of interactions within an ecosystem can keep its numbers and types of organisms relatively constant over long periods of time under stable conditions. If a modest biological or physical disturbance to an ecosystem occurs, it may return to its more or less original status (i.e., the ecosystem is resilient), as opposed to becoming a very different ecosystem. Extreme fluctuations in conditions or the size of any population, however, can challenge the functioning of ecosystems in terms of resources and habitat availability.</td><td>Stability and Change

• Much of science deals with constructing explanations of how things change and how they remain stable.</td></tr>
<tr><td>Clarifications and Content Limits</td><td colspan="3">Clarification Statements
• Examples of changes in ecosystem conditions could include modest biological or physical changes, such as moderate hunting or a seasonal flood, and extreme changes, such as volcanic eruption or sea-level rise.
• To show full comprehension of the PE, the student must demonstrate an understanding that, in a stable ecosystem, the average activity by the nutrients, decomposers, producers, primary consumers, secondary consumers, and tertiary consumers remains relatively consistent. When each of these levels has high levels of diversity, the ecosystem is stable because the group as a whole is better able to respond to pressures. However, even a healthy, diverse ecosystem is subject to extreme changes when faced with enough pressure.

Content Limits
• Assessment does not include Hardy-Weinberg equilibrium calculations.</td></tr>
<tr><td>Science Vocabulary Students Are Expected to Know</td><td colspan="3">Biosphere, biodiversity, carbon cycle, water cycle, nitrogen cycle, fluctuation, consistent, stable, equilibrium, species, emergence, extinction, niche, native, non-native, invasive, overgrazing, human impact, succession, primary succession, secondary succession.</td></tr>
<tr><td>Science Vocabulary Students Are Not Expected to Know</td><td colspan="3">Genetic drift, founder effect, Hardy-Weinberg, intermediate disturbance hypothesis, species-area curve.</td></tr>
<tr><td colspan="4">Phenomena</td></tr>
<tr><td>Context/ Phenomena</td><td colspan="3">Some example phenomena for HS-LS2-6:
• The populations of rabbits and deer in the Florida Everglades significantly decreased with the introduction of the Burmese python.
• Biodiversity of an area of the Amazon rainforest is affected differently in sustainable and non-sustainable lumber farms.</td></tr>
</table>

	• After a fire, the biodiversity of a forest immediately decreases but eventually increases. • An increase in mouse populations are observed the year after a flood but return to pre-flood numbers the following year.

This Performance Expectation and associated Evidence Statements support the following Task Demands.
Task Demands
1. Based on the provided data or information, identify the explanation that complex interactions in ecosystems maintain relatively consistent numbers and types of organisms in stable conditions, but changing conditions may result in a new ecosystem.
2. Identify and/or explain the claims, evidence, and reasoning supporting the explanation that complex interactions in ecosystems maintain relatively consistent numbers and types of organisms in stable conditions, but changing conditions may result in a new ecosystem.
3. Identify and/or describe additional relevant evidence not provided that would support or clarify the explanation of the complex interactions in ecosystems, factors that affect biodiversity, relationships between species and the environment, and changes in numbers of species and organisms in a stable or changing ecosystem.
4. Evaluate the strengths and weaknesses of a claim to explain the relationship of biodiversity and the environment in an ecosystem based on the evidence or data provided.*
5. Analyze and/or interpret evidence and its ability to support the explanation of the resiliency of an ecosystem in response to different levels of change.*
6. Provide and/or evaluate reasoning to support the explanation that an ecosystem remains relatively consistent when faced with modest disturbances, but it may experience extreme changes or fluctuations in biodiversity when faced with extreme disturbances.*

*denotes those task demands which are deemed appropriate for use in stand-alone item development

<table>
<tr><td>Performance Expectation</td><td colspan="3">HS-LS2-7
Design, evaluate, and refine a solution for reducing the impacts of human activities on the environment and biodiversity.</td></tr>
<tr><td>Dimensions</td><td>Constructing Explanations and Designing Solutions
• Design, evaluate, and refine a solution to a complex real-world problem, based on scientific knowledge, student-generated sources of evidence, prioritized criteria, and trade-off considerations.</td><td>LS2.C: Ecosystem Dynamics, Functioning, and Resilience
• Moreover, anthropogenic changes (induced by human activity) in the environment—including habitat destruction, pollution, introduction of invasive species, overexploitation, and climate change—can disrupt an ecosystem and threaten the survival of some species.

LS4.D: Biodiversity and Humans
• Biodiversity is increased by the formation of new species (speciation) and decreased by the loss of species (extinction). (secondary)

ETS1.B: Developing Possible Solutions
• When evaluating solutions, it is important to take into account a range of constraints including cost, safety, reliability, and aesthetics, and to consider social, cultural, and environmental impacts. (secondary)</td><td>Stability and Change
• Much of science deals with constructing explanations of how things change and how they remain stable.</td></tr>
<tr><td>Clarifications and Content Limits</td><td colspan="3">Clarification Statements
• Examples of human activities can include urbanization, building dams, and dissemination of invasive species.

Content Limits
• Assessment does not include physical equations describing mechanics of solutions or mechanics of engineered structures.
• <u>Students do not need to know</u>: quantitative statistical analysis, specific conditions required for failure, specifics of constructing the solution.</td></tr>
<tr><td>Science Vocabulary Students Are Expected to Know</td><td colspan="3">Carrying capacity, competition, urbanization, conversation biology, endangered species, threatened species, introduced species, overharvesting, extinction, greenhouse effect, carbon footprint</td></tr>
<tr><td>Science Vocabulary Students Are Not Expected to Know</td><td colspan="3">Laws of thermodynamics, Hardy-Weinberg equilibrium, Lotka-Volterra equations, allelopathy, density-dependent population regulation, extinction vortex, minimum viable population (MVP), effective population size, movement corridor, biodiversity hot spot, zoned reserve, critical load, biological magnification, assisted migration, sustainable development.</td></tr>
<tr><td colspan="4" align="center">Phenomena</td></tr>
<tr><td>Context/ Phenomena</td><td colspan="3">Some example of phenomena for HS-LS2-7:
• The spread of cities through urbanization has destroyed wildlife habitats across the planet.
• Air pollution from driving cars has made the air unsafe to breathe in many areas.
• Dams have led to flooding of large areas of land, destroying animal habitats.</td></tr>
</table>

	• Fishing has drastically changed marine ecosystems, removing certain predators or certain prey.
This Performance Expectation and associated Evidence Statements support the following Task Demands.	
Task Demands	
1. Articulate, describe, illustrate, or select the relationships, interactions, and/or processes to be explained. This may entail sorting relevant from irrelevant information or features.	
2. Express or complete a causal chain explaining how human activity impacts the environment. This may include indicating directions of causality in an incomplete model such as a flow chart or diagram, or completing cause-and-effect chains.	
3. Identify evidence supporting the inference of causation that is expressed in a causal chain.	
4. Use an explanation to predict the environmental outcome, given a change in the design of human technology.	
5. Describe, identify, and/or select information needed to support an explanation.	
6. Identify or describe relevant aspects of the problem that given design solutions for reducing the impacts of human activities on the environment and biodiversity, if implemented, will resolve or improve.	
7. Using given information about the effects of human activities on the environment and biodiversity, select or identify criteria against which the solution should be judged.	
8. Using given information about the effects of human activities on the environment and biodiversity, select or identify constraints that the solution must meet.	
9. Evaluate the criteria and constraints, along with trade-offs, for a proposed or given solution to resolve or improve the impact of human activities on the environment and biodiversity.	
10. Using given data, propose a potential solution to resolve or improve the impact of human activities on the environment and biodiversity.	
11. Using a simulator, test a proposed solution to resolve or improve the impact of human activities on the environment and biodiversity and evaluate the outcomes.	
12. Evaluate and/or revise a solution to resolve or improve the impact of human activities on the environment and biodiversity and evaluate the outcomes	

<table>
<tr><td>Performance Expectation</td><td colspan="3">HS-LS2-8
Evaluate the evidence for the role of group behavior on individual and species’ chances to survive and reproduce.</td></tr>
<tr><td>Dimensions</td><td>Engaging in Argument from Evidence
• Evaluate the evidence behind currently accepted explanations to determine the merits of arguments.</td><td>LS2.D: Social Interactions and Group Behavior
• Group behavior has evolved because membership can increase the chances of survival for individuals and their genetic relatives.</td><td>Cause and Effect
• Empirical evidence is required to differentiate between cause and correlation and make claims about specific causes and effects.</td></tr>
<tr><td>Clarifications and Content Limits</td><td colspan="3">Clarification Statements
• Emphasis is on: (1) distinguishing between group and individual behavior, (2) identifying evidence supporting the outcomes of group behavior, and (3) developing logical and reasonable arguments based on evidence.
• Examples of group behaviors could include flocking, schooling, herding, and cooperative behaviors such as hunting, migrating, and swarming.

Content Limits
• <u>Students do not need to know:</u>
  • How to develop or analyze computer simulations and mathematical models that emulate the flocking behavior of animals.
  • Individual genes or complex gene interactions determining individual animal behavior.</td></tr>
<tr><td>Science Vocabulary Students Are Expected to Know</td><td colspan="3">Behavioral ecology, cooperative behavior, altruism, environmental stimuli, circadian clock, communication, foraging, optimal foraging model, energy costs and benefits, competition, predator, mutual protection, packs</td></tr>
<tr><td>Science Vocabulary Students Are Not Expected to Know</td><td colspan="3">Fixed action pattern, pheromones, innate behavior, learning, imprinting, spatial learning, social learning, associative learning, problem solving, cognition, game theory, agonistic behavior, mating behavior, mating systems, parental care, mate choice, male competition for mates, reciprocal altruism, shoaling</td></tr>
<tr><td colspan="4" align="center">Phenomena</td></tr>
<tr><td>Context/ Phenomena</td><td colspan="3">Some example phenomena for HS-LS2-8:
• Several hundred naked mole rats are observed living together in a colony. However, only one large naked mole rat is observed reproducing, while the others in the colony bring her food.
• A worker bee is observed flying away from its colony. Upon returning many other worker bees crowd around him while he moves in a distinct pattern.
• A lioness charges toward a large herd of galloping zebra, but then stops and runs away in the opposite direction.
• A certain species of short-horned grasshoppers changes color, band together, and fly over several square kilometers over a period of a few weeks.</td></tr>
<tr><td colspan="4" align="center">This Performance Expectation and associated Evidence Statements support the following Task Demands.</td></tr>
<tr><td colspan="4" align="center">Task Demands</td></tr>
</table>

1. Based on the provided data, identify, describe, or construct a claim regarding how specific group behavior(s) can increase an individual's or species' chances of surviving and reproducing.
2. Sort inferences about the effect of specific group behaviors on an individual's and species' chances to survive and reproduce into those that are supported by the data, contradicted by the data, outliers in the data, or neither, or some similar classification.
3. Identify patterns of information/evidence in the data that support correlative/causative inferences about the effect of specific group behaviors on an individual's and species' chances to survive and reproduce.*
4. Construct an argument using scientific reasoning, drawing on credible evidence to explain the effect of specific group behaviors on an individual's and species' chances to survive and reproduce.
5. Identify additional evidence that would help clarify, support, or contradict a claim or causal argument regarding the effect of specific group behaviors on an individual's and species' chances to survive and reproduce.
6. Identify, summarize, or organize given data or other information to support or refute a claim regarding the effect of specific group behaviors on an individual's and species' chances to survive and reproduce.**

*denotes those task demands which are deemed appropriate for use in stand-alone item development
**TD6 – summarize is the emphasis here. Avoid identify and organize.

<table>
<tr><td>Performance Expectation</td><td colspan="3">HS-LS3-1
Ask questions to clarify relationships about the role of DNA and chromosomes in coding the instructions for characteristic traits passed from parents to offspring.</td></tr>
<tr><td>Dimensions</td><td>Asking Questions and Defining Problems
• Ask questions that arise from examining models or a theory to clarify relationships.</td><td>LS1.A: Structure and Function
• All cells contain genetic information in the form of DNA molecules. Genes are regions in the DNA that contain the instructions that code for the formation of proteins. (secondary)

LS3.A: Inheritance of Traits
• Each chromosome consists of a single very long DNA molecule, and each gene on the chromosome is a particular segment of that DNA. The instructions for forming species’ characteristics are carried in DNA. All cells in an organism have the same genetic content, but the genes used (expressed) by the cell may be regulated in different ways. Not all DNA codes for a protein; some segments of DNA are involved in regulatory or structural functions, and some have no as-yet known function.</td><td>Cause and Effect
• Empirical evidence is required to differentiate between cause and correlation and to make claims about specific causes and effects.</td></tr>
<tr><td>Clarifications and Content Limits</td><td colspan="3">Clarification Statements:
• At this level, the study of inheritance is restricted to Mendelian genetics, including dominance, codominance, incomplete dominance, and sex-linked traits.
• Focus is on expression of traits on the organism level and should not be restricted to protein production.

Content Limits:
• Assessment does not include the phases of meiosis or the biochemical mechanism of specific steps in the process.
• Assessment does not include mutations or species-level genetic variation including Hardy-Weinberg equilibrium.</td></tr>
<tr><td>Science Vocabulary Students Are Expected to Know</td><td colspan="3">Genome, zygote, fertilization, dominant, recessive, codominance, incomplete dominance, sex-linked, allele, sequencing, pedigree, parent generation, F1, F2, haploid, diploid, replication.</td></tr>
<tr><td>Science Vocabulary Students Are Not Expected to Know</td><td colspan="3">Epigenetics, interphase, prophase, metaphase, anaphase, telophase, cytokinesis, epistasis.</td></tr>
<tr><td colspan="4">Phenomena</td></tr>
<tr><td>Context/ Phenomena</td><td colspan="3">Some example phenomena for HS-LS3-1:
• DNA sequencing shows that all people have the gene for lactase production, but only about 30% of adults can digest milk.
• Polydactyl tabby cat Jake holds the world record for most toes, with seven toes on each paw.</td></tr>
</table>

	• *E. coli* bacteria are healthful in mammalian intestines but makes mammals sick when ingested. • *E. coli* bacteria are used to produce human insulin.
This Performance Expectation and associated Evidence Statements support the following Task Demands.	
Task Demands	
8. Identify or construct an empirically testable question based on the phenomenon that could lead to design of an experiment or model to define the relationships between the role of DNA and/or chromosomes in the inheritance of traits.*	
9. Assemble or complete, from a collection of potential model components, an illustration, or pedigree that is capable of representing the role of genetic material in coding the instructions for inheritance.*	
10. Construct a question that arises from examining a model or theory to clarify the connections between DNA/chromosomes and inheritance of traits.*	
11. Make predictions about the pattern of inheritance based on a model derived from the empirically testable question. Predictions can be made by manipulating model components, completing illustrations, or selecting from lists with distractors.	
12. Assemble or complete a flow chart describing the cause and effect relationships between genetic material and the characteristic traits passed from parents to offspring.	

*denotes those task demands which are deemed appropriate for use in stand-alone item development

<table>
<tr><td>Performance Expectation</td><td colspan="3">HS-LS3-2
Make and defend a claim based on evidence that inheritable genetic variations may result from: (1) new genetic combinations through meiosis, (2) viable errors occurring during replication, and/or (3) mutations caused by environmental factors.</td></tr>
<tr><td>Dimensions</td><td>Engaging in Argument from Evidence
• Make and defend a claim based on evidence about the natural world that reflects scientific knowledge and student-generated knowledge.</td><td>LS3.B: Variation of Traits
▪ In sexual reproduction, chromosomes can sometimes swap sections during the process of meiosis (cell division), thereby creating new genetic combinations and thus more genetic variation. Although DNA replication is tightly regulated and remarkably accurate, errors do occur and result in mutations, which are also a source of genetic variation. Environmental factors can also cause mutations in genes, and viable mutations are inherited.
▪ Environmental factors also affect expression of traits, and, hence, they affect the probability of occurrences of traits in a population. Thus, the variation and distribution of traits observed depends on both genetic and environmental factors.</td><td>Cause and Effect
• Empirical evidence is required to differentiate between cause and correlation, and to make claims about specific causes and effects.</td></tr>
<tr><td>Clarifications and Content Limits</td><td colspan="3">Clarification Statements
• Emphasis is on using data to support arguments for the way variation occurs.
• Inheritable traits should be traits that can be passed down through more than one generation.
• Inheritable traits for this PE do not include dominant/recessive traits.
• Examples of evidence for new genetic combinations and viable errors can include:
    • karyotype comparison between parents and children;
    • DNA sequence comparison.

Content Limits
• Assessment does not include assessing meiosis or the biochemical mechanism of specific steps in the process.
• <u>Students do not need to know</u>: bioinformatics, specific genetic disorders.</td></tr>
<tr><td>Science Vocabulary Students Are Expected to Know</td><td colspan="3">Amino acid, DNA, enzyme, protein synthesis, chromosome, egg, egg cell, sperm, sperm cell, dominant trait, recessive trait, recombination, sex cell, sex chromosome, sex-linked trait, meiosis, mutation, advantageous, expression, base pairs, genome, UV radiation, triplet codon, insertion, deletion, frameshift, substitution, somatic, epigenetic.</td></tr>
<tr><td>Science Vocabulary Students Are Not Expected to Know</td><td colspan="3">Polyploidy, single nucleotide polymorphisms (SNPs), conjugation, DNA polymerase, mutagenic, chromosomal translocation, missense, nonsense, nongenic region, tautomerism, depurination, deamination, slipped-strand mispairing, Sheik disorder, prion, epidemiology.</td></tr>
<tr><td colspan="4">Phenomena</td></tr>
<tr><td>Context/ Phenomena</td><td colspan="3">Some example phenomena for HS-LS3-2:
• Due to pesticide residue, frogs have extra, non-functioning, limbs.
• Most chickens have feathers that lay flat against their bodies. In one family of chickens, 50% of offspring have feathers that curl away from their bodies.</td></tr>
</table>

	• A single gene mutation accounts for the blue color of irises in over 99.5% of people with blue eyes. • One sunflower growing in a field has a wide, flat stem and an unusual number of leaves. The next year, several sunflowers in the field share these traits.
This Performance Expectation and associated Evidence Statements support the following Task Demands.	
Task Demands	
1. Based on the provided data, make or construct a claim regarding inheritable genetic variations that may result from: 1) new genetic combinations through meiosis, 2) viable errors occurring during replication, and/or 3) mutations caused by environmental factors. This *does not* include selecting a claim from a list.	
2. Sort inferences about inheritable genetic variation into those that are supported by the data, contradicted by the data, outliers in the data, or none of these—or some similar classification.	
3. Identify patterns of information/evidence in the data that support correlative/causative inferences about inheritable genetic variation.	
4. Construct an argument using scientific reasoning that draws on credible evidence to explain how inheritable genetic variations may result from: (1) new genetic combinations through meiosis, (2) viable errors occurring during replication, and/or (3) mutations caused by environmental factors. (Hand scored CR)	
5. Identify additional evidence that would help clarify, support, or contradict a claim or causal argument.	
6. Identify, describe, and/or construct alternate explanations or claims, and cite the data needed to distinguish among them.	
7. Predict outcomes of genetic variations, given the cause-and-effect relationships of inheritance.	

Performance Expectation	**HS-LS3-3** Apply concepts of statistics and probability to explain the variation and distribution of expressed traits in a population.		
Dimensions	**Analyzing and Interpreting Data** • Apply concepts of statistics and probability (including determining function fits to data, slope, intercept, and correlation coefficient for linear fits) to scientific and engineering questions and problems, using digital tools when feasible.	**LS3.B Variation of Traits** • Environmental factors also affect expression of traits, and hence affect the probability of occurrences of traits in a population. Thus, the variation and distribution of traits observed depends on both genetic and environmental factors.	**Scale, Proportion and Quantity** • Algebraic thinking is used to examine scientific data and predict the effect of a change in one variable on another (e.g., linear growth vs. exponential growth).
Clarifications and Content Limits	**Clarification Statements** • Emphasis is on the use of mathematics to describe the probability of traits as it relates to genetic and environmental factors in the expression of traits. • Sensitivity and precaution should be used around the use of both lethal recessive and dominant human traits (i.e., Huntington’s, achondroplasia, Tay-Sachs, cystic fibrosis). **Content Limits** • Assessment is limited to basic statistical and graphical analysis. • Assessment does not include Hardy-Weinberg calculations ($p^2 + 2pq + q^2 = 1$ or $p + q = 1$). • <u>Students do not need to know:</u> pleiotropy, meiosis, specific names of genetic disorders.		
Science Vocabulary Students are Expected to Know	Gene, allele, dominant, recessive, homozygous, heterozygous, phenotype, genotype, P generation, F_1 generation, F_2 generation, complete dominance, incomplete dominance, codominance, pedigree, carrier, fertilization, sex linked traits, gamete, Mendelian genetics, zygote, haploid, diploid, epistasis.		
Science Vocabulary Students are Not Expected to Know	Test-cross, monohybrid, dihybrid, law of independent assortment, law of segregation, pleiotropy, norm of reaction, multifactorial, Barr Body, genetic recombination, latent allele.		
Phenomena			
Context/ Phenomena	Some example phenomena for HS-LS3-3: • O Positive is the most common blood type. Not all ethnic groups have the same mix of these blood types. Hispanic people, for example, have a relatively high number of O’s, while Asian people have a relatively high number of B’s. • Hydrangea flowers of the same genetic variety range in color from blue-violet to pink, with the shade and intensity of color depending on the acidity and aluminum content of the soil. • Most humans were born with five fingers on each hand, yet the polydactyl trait (having more than five fingers on each hand) is the dominant trait. • When a red rose is crossed with a white rose, all pink roses are produced.		
This Performance Expectation and associated Evidence Statements support the following Task Demands.			
Task Demands			

1. Describe data or patterns/relationships in given data that support (or refute) an explanation for the change in trait frequency or magnitude in a population, due to both genetic and environmental factors.*
2. Make predictions about the trait frequency or distribution in a population due to the presence/absence or addition/removal of both genetic and environmental factors.*
3. Organize and/or arrange (e.g., using illustrations and/or labels) data, or summarize data to provide evidence for an explanation of the relationship between a trait's occurrence in a population and genetic and environmental factors.
4. Analyze, evaluate, estimate, calculate, and/or construct an equation for the statistical mean and/or the standard deviation, to describe the change in the distribution of a trait in a population over time, due to genetic and environmental factors.*
5. Identify statistical anomalies or outliers for a trait or a population that are outside the expected range (norm reaction), which may or may not be quickly removed due to genetic and environmental factors.

*denotes those task demands which are deemed appropriate for use in stand-alone item development

<table>
<tr><td>Performance Expectation</td><td colspan="3">HS-LS4-1
Communicate scientific information that common ancestry and biological evolution are supported by multiple lines of empirical evidence.</td></tr>
<tr><td>Dimensions</td><td>Obtaining, Evaluating, and Communicating Information
• Communicate scientific information (e.g. about phenomena and/or the process of development and the design and performance of a proposed process or system) in multiple formats (including orally, graphically, textually, and mathematically).</td><td>LS4.A: Evidence of Common Ancestry and Diversity
• Genetic information, like the fossil record, provides evidence of evolution. DNA sequences vary among species, but there are many overlaps; in fact, the ongoing branching that produces multiple lines of descent can be inferred by comparing the DNA sequences of different organisms. Such information is also derivable from the similarities and differences in amino acid sequences and from anatomical and embryological evidence.</td><td>Patterns
• Different patterns may be observed at each of the scales at which a system is studied and can provide evidence for causality in explanations of phenomena.</td></tr>
<tr><td>Clarifications and Content Limits</td><td colspan="3">Clarification Statements
• Emphasis is on a conceptual understanding of the role each line of evidence has relating to common ancestry and biological evolution.
• Examples of evidence could include similarities in DNA sequences, anatomical structures, and order of appearance of structures in embryological development.

Content Limits
• <u>Students do not need to know</u>: specific genetic mutations, specific genetic disorders, specific proteins, Occam’s razor (maximum parsimony), formation of orthologous and paralogous genes, molecular clock, Neutral theory.</td></tr>
<tr><td>Science Vocabulary Students are Expected to Know</td><td colspan="3">Amino acids, cladogram, comparative anatomy, DNA sequencing, electrophoresis, embryology, evolution, fossil record, gene flow, genetic drift, mutation, natural selection, nucleotides, sedimentary layers, species, descent with modification, homologous structures, evolutionary tree, analogous structures.</td></tr>
<tr><td>Science Vocabulary Students are Not Expected to Know</td><td colspan="3">Phylogenetic, phylogeny, phylogenetic tree, taxonomy, cladistics, vestigial structures, convergent evolution, analogous, endemic, phylocode, sister taxa, basal taxon, polytomy, homoplasy, molecular systematics, monophyletic, parphyletic, polyphyletic, maximum parsimony, orthologous genes, paralogous genes, horizontal gene transfer.</td></tr>
<tr><td colspan="4">Phenomena</td></tr>
<tr><td>Context/ Phenomena</td><td colspan="3">Some example phenomena for HS-LS4-1:
• Red pandas look a bit like bears and a bit like raccoons. Task Statement: Provide evidence about whether red pandas are better classified as raccoons or bears. Stimulus material might include pictures, DNA information, embryological information, and homologous structures.
• Hermit crabs live in shells, like oysters, but look like crabs. Provide evidence classifying hermit crabs either as mollusks (like oysters) or arachnids (like crabs).
• Crawfish look just like lobster, but smaller. Which came first, the lobster or the crawfish?
• Fossil records of an extinct hooved animal show a thickened knob of bone in its middle ear. This structure is also found in modern whales and helps them hear underwater.</td></tr>
</table>

This Performance Expectation and associated Evidence Statements support the following Task Demands.
Task Demands
1. Analyze and interpret scientific evidence from multiple scientific/technical sources including text, diagrams, charts, symbols, mathematical representations that support common ancestry among organisms and/or biological evolution.*
2. Evaluate the validity/relevance/reliability of scientific evidence about biological evolution.
3. Identify relationships or patterns in scientific evidence at macroscopic and/or microscopic scales.*
4. Describe the specific evidence needed to support an explanation about how organisms share a common ancestor.
5. Synthesize an explanation that incorporates the scientific evidence from multiple sources.

*denotes those task demands which are deemed appropriate for use in stand-alone item development

<table>
<tr><td>Performance Expectation</td><td colspan="3">HS-LS4-2
Construct an explanation based on evidence that the process of evolution primarily results from four factors: (1) the potential for a species to increase in number, (2) the heritable genetic variation of individuals in a species due to mutation and sexual reproduction, (3) competition for limited resources, and (4) the proliferation of those organisms that are better able to survive and reproduce in the environment.</td></tr>
<tr><td>Dimensions</td><td>Constructing Explanations and Designing Solutions
• Construct an explanation based on valid and reliable evidence obtained from a variety of sources (including students’ own investigations, models, theories, simulations, and peer review) and the assumption that theories and laws that describe the natural world operate today as they did in the past and will continue to do so in the future.</td><td>LS4.B: Natural Selection
▪ Natural selection occurs only if there is both 1) variation in the genetic information between organisms in a population and 2) variation in the expression of that genetic information—that is, trait variation—that leads to differences in performance among individuals.

LS4.C: Adaptation
• Evolution is a consequence of the interaction of four factors: 1) the potential for a species to increase in number, 2) the genetic variation of individuals in a species due to mutation and sexual reproduction, 3) competition for an environment’s limited supply of the resources that individuals need in order to survive and reproduce, and 4) the ensuing proliferation of those organisms that are better able to survive and reproduce in that environment.</td><td>Cause and Effect
• Empirical evidence is required to differentiate between cause and correlation and to make claims about specific causes and effects.</td></tr>
<tr><td>Clarifications and Content Limits</td><td colspan="3">Clarification Statements
• Emphasis is on using evidence to explain the influence each of the four factors has on the number of organisms, behaviors, morphology, or physiology in terms of ability to compete for limited resources and subsequent survival of individuals and adaptation of species.
• Examples of evidence could include mathematical models such as simple distribution graphs and proportional reasoning.

Content Limits
• Assessment does not include other mechanisms of evolution, such as genetic drift, gene flow through migration, and co-evolution.
• Students do not need to know: Hardy-Weinberg equation.</td></tr>
<tr><td>Science Vocabulary Students Are Expected to Know</td><td colspan="3">Beneficial change, detrimental change, distribution, emergence, gene frequency, biotic, abiotic, advantageous, diverge, proliferation, bottleneck effect, island effect, geographic isolation, founder effect, recombination.</td></tr>
<tr><td>Science Vocabulary Students Are Not Expected to Know</td><td colspan="3">Hardy-Weinberg equilibrium, biotechnology, relative fitness, directional selection, disruptive selection, stabilizing selection, heterozygote advantage, frequency-dependent selection, prezygotic barriers, postzygotic barriers.</td></tr>
<tr><td colspan="4">Phenomena</td></tr>
<tr><td>Context/</td><td colspan="3">Some example phenomena for HS-LS4-2:</td></tr>
</table>

Phenomena	• Cane toads introduced to Australia in the 1930s have evolved to be bigger, more active, and have longer legs. • In the late 1990s, a resurgence of bedbug outbreaks began. Bedbugs are now much harder to kill with thick, waxy exoskeletons, faster metabolism, and mutations to block certain insecticides. • Skinks living in cooler regions give live birth, while those living in warm coastal areas lay eggs. • A butterfly parasite found on the Samoan Islands destroyed the male embryos of blue moon butterflies, decreasing the male population to only 1%. After a year, males had rebounded to 40% of the population.

This Performance Expectation and associated Evidence Statements support the following Task Demands.
Task Demands
1. Describe the cause-and-effect relationship between: (1) the potential for a species to increase in number, (2) the heritable genetic variation of individuals in a species due to mutation and sexual reproduction, (3) competition for limited resources, and (4) the proliferation of those organisms that are better able to survive and reproduce in the environment, and change in species over time. This may include indicating directions of causality in a model or completing cause-and-effect chains.
2. Describe, identify, or select evidence supporting or contradicting a claim about the role of (1) the potential for a species to increase in number, (2) the heritable genetic variation of individuals in a species due to mutation and sexual reproduction, (3) competition for limited resources, and (4) the proliferation of those organisms that are better able to survive and reproduce in the environment in causing the phenomenon. The evidence may be evidence generated by the students in the simulation or selected from provided data.
3. Given an appropriate explanation for a phenomenon, predict the effects of subsequent changes in environmental conditions on the population.
4. Use evidence to construct an explanation of the changes in species over time as a result of (1) the potential for a species to increase in number, (2) the heritable genetic variation of individuals in a species due to mutation and sexual reproduction, (3) competition for limited resources, and (4) the proliferation of those organisms that are better able to survive and reproduce in the environment.***(SEP/DCI/CCC)**
5. Identify and justify additional pieces of evidence that would help distinguish between competing hypotheses for the changes in species over time.

*denotes those task demands which are deemed appropriate for use in stand-alone item development

<table>
<tr><td>Performance Expectation</td><td colspan="3">HS-LS4-3
Apply concepts of statistics and probability to support explanations that organisms with an advantageous heritable trait tend to increase in proportion to organisms lacking this trait.</td></tr>
<tr><td>Dimensions</td><td>Analyzing and Interpreting Data
• Apply concepts of statistics and probability (including determining function fits to data, slope, intercept, and correlation coefficient for linear fits) to scientific and engineering questions and problems, using digital tools when feasible.</td><td>LS4.B: Natural Selection
• Natural selection occurs only if there is both (1) variation in the genetic information between organisms in a population and (2) variation in the expression of that genetic information—that is, trait variation —that leads to differences in performance among individuals.
• The traits that positively affect survival are more likely to be reproduced, and thus are more common in the population.

LS4.C: Adaptation
• Natural selection leads to adaptation that is, to a population dominated by organisms that are anatomically, behaviorally, and physiologically well suited to survive and reproduce in a specific environment. The differential survival and reproduction of organisms in a population that have an advantageous heritable trait lead to an increase in the proportion of individuals in future generations that have the trait and to a decrease in the proportion of individuals that do not.
• Adaptation also means that the distribution of traits in a population can change when conditions change.</td><td>Patterns
• Different patterns may be observed at each of the scales at which a system is studied and can provide evidence for causality in explanations of phenomena.</td></tr>
<tr><td>Clarifications and Content Limits</td><td colspan="3">Clarification Statements
• Emphasis is on analyzing shifts in numerical distribution of traits and using these shifts as evidence to support explanations.

Content Limits
• Assessment is limited to basic statistical and graphical analysis. Assessment does not include allele frequency calculations.
• <u>Students do not need to know</u>: sexual selection, kin selection, artificial selection, frequency-dependent selection.</td></tr>
<tr><td>Science Vocabulary Students are Expected to Know</td><td colspan="3">Fitness, gene, allele, directional selection, diversifying (disruptional selection), stabilizing selection, standard deviation, vestigial structure.</td></tr>
<tr><td>Science Vocabulary Students are Not Expected to Know</td><td colspan="3">Hemizygous, aneuploidy, intragenomic conflict, sexual dimorphism, balanced polymorphism, apostatic selection.</td></tr>
<tr><td colspan="4" align="center">Phenomena</td></tr>
<tr><td>Context/ Phenomena</td><td colspan="3">Example phenomena for HS-LS4-3:</td></tr>
</table>

	• Green Treefrogs (*Hyla versicolor*) are abundant in the wetlands of Florida where no Gray Treefrogs (*Hyla cinerea*) are observed. In the wooded areas of New York, only Gray Treefrogs are observed. • In the Amazon rainforest, a kapok trees (*Ceiba pentandra*) measures 200 feet in height, approximately 30 feet above the rest of the canopy. • A school of mummichog fish (*Fundulus heteroclitus*) is found in the 6°C waters of the Chesapeake Bay. These fish are normally found in warmer climates, like the 21°C waters of Kings Bay, Georgia. • A population of the fish *Poecilia mexicana* lives in the murky hydrogen-sulfide (H2S)-rich waters in southern Mexico that would kill the same species of fish living in clear freshwaters only 10 km away.

This Performance Expectation and associated Evidence Statements support the following Task Demands.
Task Demands
1. Describe or identify patterns or relationships in given data that support (or refute) an explanation for the change in trait frequency or magnitude in a population due to natural selection/selection pressure(s).*
2. Make predictions about the trait frequency or distribution in a population due to the presence/absence or addition/removal of selection pressure(s) in the environment (including Hardy-Weinberg-based predictions about changes in allele/trait frequency/magnitude NOT based on calculations).*
3. Organize and/or arrange (e.g., using illustrations and/or labels) data, or summarize data to provide evidence for an explanation of the effect of selection on a population.
4. Analyze, evaluate, estimate, calculate, and/or construct an equation to describe the change in the distribution of a trait in a population over time due to selection pressure(s).
5. Identify statistical anomalies or outliers for a trait or a population that are outside the expected range (for example, Joe DiMaggio's hitting streak, tossing 10 consecutive heads on a fair coin, etc.) which may or may not be quickly removed due to selection pressure.
6. Use statistical analysis to calculate changes in traits in a population over time to provide evidence for an explanation of the relationship between a trait's occurrence and its prevalence in the population at different points in time.
7. Identify explanations for a change in a traits frequency and/or distribution in a population over time that can be supported by patterns or relationships in data.

*denotes those task demands which are deemed appropriate for use in stand-alone item development

<table>
<tr><td>Performance Expectation</td><td colspan="3">HS-LS4-4
Construct an explanation based on evidence for how natural selection leads to adaptation of populations.</td></tr>
<tr><td>Dimensions</td><td>Constructing Explanations and Designing Solutions
• Construct an explanation based on valid and reliable evidence obtained from a variety of sources (including students’ own investigations, models, theories, simulations, peer review) and the assumption that theories and laws that describe the natural world operate today as they did in the past and will continue to do so in the future.</td><td>LS4.C: Adaptation
• Natural selection leads to adaptation; that is, to a population dominated by organisms that are anatomically, behaviorally, and physiologically well suited to survive and reproduce in a specific environment. The differential survival and reproduction of organisms in a population that has an advantageous, heritable trait leads to an increase in the proportion of individuals in future generations that have the trait and to a decrease in the proportion of individuals that do not.</td><td>Cause and Effect
• Empirical evidence is required to differentiate between cause and correlation and make claims about specific causes and effects.</td></tr>
<tr><td>Clarifications and Content Limits</td><td colspan="3">Clarification Statement
• Emphasis is on using data to provide evidence for how specific biotic and abiotic differences in ecosystems (such as ranges of seasonal temperature, long-term climate change, acidity, light, geographic barriers, or evolution of other organisms) contribute to a change in gene frequency over time, leading to adaptation of populations.

Content Limits
• Assessment does not include the Hardy-Weinberg equation.</td></tr>
<tr><td>Science Vocabulary Students Are Expected to Know</td><td colspan="3">Beneficial change, detrimental change, distribution, emergence, gene frequency, gene, biotic, abiotic, advantageous, diverge, proliferation, sexual reproduction, bottleneck effect, island effect, geographic isolation, gene flow, genetic drift, founder effect.</td></tr>
<tr><td>Science Vocabulary Students Are Not Expected to Know</td><td colspan="3">Hardy Weinberg Equilibrium, biotechnology, relative fitness, directional selection, disruptive selection, stabilizing selection, heterozygote advantage, frequency-dependent selection, prezygotic barriers, postzygotic barriers.</td></tr>
<tr><td colspan="4">Phenomena</td></tr>
<tr><td>Context/ Phenomena</td><td colspan="3">Some example phenomena for HS-LS4-4:
• Following a four-year drought in California, field mustard plants are found to flower earlier in the season.
• A new antibiotic is discovered. Within ten years, many bacterial diseases that were previously treated by the antibiotic no longer respond to treatment (e.g., MRSA).
• A small population of Italian wall lizards that feed mainly on insects is introduced to a neighboring island. After several decades, the lizards are found to have thrived and heavily populated the island, and their diet is now mostly vegetation.
• Following climatic changes, the European Great Tit bird begins laying eggs earlier in the spring.</td></tr>
<tr><td colspan="4">This Performance Expectation and associated Evidence Statements support the following Task Demands.</td></tr>
</table>

Task Demands
1. Organize or summarize the given data or evidence of population characteristics, environmental characteristics, and/or the relationships between them.
2. Generate or construct graphs or tables of data to highlight patterns within the given data.
3. Describe the cause and effect relationship between natural selection and adaptation using evidence. This may include assembling descriptions from illustrations or lists of options and distractors, or indicating directions of causality in a model or completing cause and effect chains.
4. Describe, identify, or select evidence supporting or contradicting a claim about the role of adaptation in causing the phenomenon. The evidence may be generated by the students in a simulation.
5. Given an appropriate explanation for a phenomenon, predict the effects of subsequent changes in environmental conditions on the population.
6. Use evidence to construct an explanation of the adaptation of a species through natural selection. Evidence can be described, identified, or selected/assembled from lists with distractors. Explanations can be written, assembled by manipulating the components of a flow chart or model, or assembled from lists of options that include distractors. Options and distractors should not be single words or short phrases; rather, they should be complete thoughts that, when correctly emplaced within a sentence or paragraph, work to provide evidence of a coherent train of thought.*
7. Identify and justify additional pieces of evidence that would help distinguish among competing hypotheses.

*denotes those task demands which are deemed appropriate for use in stand-alone item development

<table>
<tr><td>Performance Expectation</td><td colspan="3">HS-LS4-5
Evaluate the evidence supporting claims that changes in environmental conditions may result in: (1) increases in the number of individuals of some species, (2) the emergence of new species over time, and (3) the extinction of other species.</td></tr>
<tr><td>Dimensions</td><td>Engaging in Argument from Evidence
• Evaluate the evidence behind currently accepted explanations or solutions to determine the merits of arguments.</td><td>LS4.C: Adaptation
• Changes in the physical environment, whether naturally occurring or human induced, have thus contributed to the expansion of some species, the emergence of new distinct species as populations diverge under different conditions, and the decline—and sometimes extinction—of some species.
• Species become extinct because they can no longer survive and reproduce in their altered environment. If members cannot adjust to change that is too fast or drastic, the opportunity for the species’ evolution is lost.</td><td>Cause and Effect
• Empirical evidence is required to differentiate between cause and correlation and make claims about specific causes and effects.</td></tr>
<tr><td>Clarifications and Content Limits</td><td colspan="3">Clarification Statements
• Emphasis is on determining cause and effect relationships for how changes to the environment such as deforestation, fishing, application of fertilizers, drought, flood, and the rate of change of the environment affect distribution or disappearance of traits in species.

Content Limits
• <u>Students do not need to know</u>: Hardy Weinberg Equation.</td></tr>
<tr><td>Science Vocabulary Students Are Expected to Know</td><td colspan="3">Beneficial change, detrimental change, distribution, emergence, gene frequency, biotic, abiotic, advantageous, diverge, mutation, proliferation, bottleneck effect, island effect, geographic isolation, founder effect, recombination, microevolution, gene flow, speciation, hybrid</td></tr>
<tr><td>Science Vocabulary Students Are Not Expected to Know</td><td colspan="3">Biotechnology, relative fitness, directional selection, disruptive selection, stabilizing selection, heterozygote advantage, frequency dependent selection, prezygotic barriers, postzygotic barriers, average heterozygosity, cline, sexual selection, sexual dimorphism, intrasexual selection, intersexual selection, neutral variation, balancing selection</td></tr>
<tr><td colspan="4" align="center">Phenomena</td></tr>
<tr><td>Context/ Phenomena</td><td colspan="3">Some example phenomena for HS-LS4-5:
• PCB pollution in the Hudson River wiped out many fish species, but the Atlantic tomcod thrives there (results 1 and 3).
• The population of Greater Prairie Chickens in Illinois decreased from millions of birds in the 1800s to fewer than 50 birds in 1993 (result 3).
• In 1681 the dodo bird went extinct due to hunting and introduction of invasive species (result 3).
• In 1988, the Orange-Spotted Filefish went extinct in response to warmer ocean temperatures (result 3).</td></tr>
<tr><td colspan="4" align="center">This Performance Expectation and associated Evidence Statements support the following Task Demands.</td></tr>
<tr><td colspan="4" align="center">Task Demands</td></tr>
</table>

1. Based on the provided data, identify, describe, or construct a claim regarding the effect of changes to the environment on the (1) increases in the number of individuals of some species, (2) the emergence of new species over time, and (3) the extinction of other species.
2. Sort inferences about the effect of changes to the environment on (1) increases in the number of individuals of some species, (2) the emergence of new species over time, and (3) the extinction of other species into those that are supported by the data, contradicted by the data, outliers in the data, or neither, or some similar classification.*
3. Identify patterns of information/evidence in the data that support correlative/causative inferences about the effect of changes to the environment on the (1) increases in the number of individuals of some species, (2) the emergence of new species over time, and (3) the extinction of other species.*
4. Construct an argument and/or explanation using scientific reasoning drawing on credible evidence to explain the effect of changes to the environment on the (1) increases in the number of individuals of some species, (2) the emergence of new species over time, and (3) the extinction of other species.
5. Identify additional evidence that would help clarify, support, or contradict a claim or causal argument regarding the effect of changes to the environment on the (1) increases in the number of individuals of some species, (2) the emergence of new species over time, and (3) the extinction of other species.
6. Identify, summarize, or organize given data or other information to support or refute a claim regarding the effect of changes to the environment on (1) increases in the number of individuals of some species, (2) the emergence of new species over time, and (3) the extinction of other species.*

*denotes those task demands which are deemed appropriate for use in stand-alone item development

<table>
<tr><td>Performance Expectation</td><td colspan="3">HS-LS4-6
Create or revise a simulation to test a solution to mitigate adverse impacts of human activity on biodiversity.</td></tr>
<tr><td>Dimensions</td><td>Using Mathematics and Computational Thinking
• Create or revise a simulation of a phenomenon, designed device, process, or system.</td><td>LS4.C: Adaptation
• Changes in the physical environment, whether naturally occurring or human induced, have thus contributed to the expansion of some species, the emergence of new distinct species as populations diverge under different conditions, and the decline—and sometimes the extinction—of some species.

LS4.D: Biodiversity and Humans
• Humans depend on the living world for the resources and other benefits provided by biodiversity. But human activity is also having adverse impacts on biodiversity through overpopulation, overexploitation, habitat destruction, pollution, introduction of invasive species, and climate change. Thus, sustaining biodiversity so that ecosystem functioning and productivity are maintained is essential to supporting and enhancing life on Earth. Sustaining biodiversity also aids humanity by preserving landscapes of recreational or inspirational value.

ETS1.B: Developing Possible Solutions
• When evaluating solutions, it is important to take into account a range of constraints, including cost, safety, reliability, and aesthetics, and to consider social, cultural, and environmental impacts (secondary).
• Both physical models and computers can be used in various ways to aid in the engineering design process. Computers are useful for a variety of purposes, such as running simulations to test different ways of solving a problem or to see which one is most efficient or economical, and in making a persuasive presentation to a client about how a given design will meet his or her needs (secondary).</td><td>Cause and Effect
• Empirical evidence is required to differentiate between cause and correlation and make claims about specific causes and effects.</td></tr>
<tr><td>Clarifications and Content Limits</td><td colspan="3">Clarification Statements
• Emphasis is on designing solutions for a proposed problem related to threatened or endangered species, or to genetic variation of organisms for multiple species.
• The simulation should model the effect of human activity and provide quantitative information about the effect of solutions on threatened or endangered species or to genetic variation within a species.

Content Limits
• <u>Students do not need to know</u>: Calculus/advanced mathematics (e.g., exponential growth and decay)</td></tr>
</table>

Science Vocabulary Students Are Expected to Know	Anthropogenic, efficient, overexploitation, urbanization, acidification, deforestation, concentration, radiation, greenhouse gas, surface runoff, civilization, consumption, mass wasting, urban development, per-capita, degradation, pollutant, best practice, cost-benefit, extract, regulation
Science Vocabulary Students Are Not Expected to Know	Oligotrophic and eutrophic lakes/eutrophication, littoral zone, exponential population growth, logistic population growth, ecological footprint, ecosystem services, extinction vortex, minimum viable population, effective population size, critical load.
Phenomena	
Context/ Phenomena	Some example phenomena for HS-LS4-6: • The habitat of the Florida Panther is only 5% of its former range, causing the species to become endangered. • The café marron plant is critically endangered due to massive habitat destruction on the Island of Rodrigues in the Indian Ocean, as a result of deforestation for agricultural use. • The population of Atlantic Bluefin Tuna has declined by more than 80% since 1970 due to overfishing. • In the past 120 years, about eighty percent of suitable orangutan habitat in Indonesia has been lost from expansion of oil palm plantations. At the same time, the estimated number of orangutans on Borneo, an island in Indonesia, has declined from about 230,000 to about 54,000.
This Performance Expectation and associated Evidence Statements support the following Task Demands.	
Task Demands	
1. Use data to calculate or estimate the effect of a solution on mitigating the adverse impacts of human activity on biodiversity.	
2. Illustrate, graph, or identify features or data that can be used to determine how effective a solution is for mitigating the adverse impacts of human activity on biodiversity.	
3. Estimate or infer the properties or relationships that lead to mitigation of the adverse impacts of human activity on biodiversity, based on data.	
4. Compile the data needed for an inference about the impacts of human activity on biodiversity. This can include sorting out the relevant data from the given information.	
5. Using given information, select or identify the criteria against which the solution should be judged.	
6. Using a simulator, test a proposed solution and evaluate the outcomes; may include proposing modifications to the solution.*	

*In order to satisfy this PE, the student must use a simulator. Therefore, this task demand must always be used.

Performance Expectation	**HS-ESS1-1** Develop a model based on evidence to illustrate the life span of the sun and the role of nuclear fusion in the sun's core to release energy that eventually reaches Earth in the form of radiation.		
Dimensions	**Developing and using models** • Develop a model based on evidence to illustrate the relationships between systems or between components of a system.	**ESS1.A: The Universe and Its Stars** • The star called the sun is changing and will burn out over a lifespan of approximately 10 billion years. **PS3.D: Energy in Chemical Processes and Everyday Life** • Nuclear Fusion processes in the center of the sun release the energy that ultimately reaches Earth as radiation. *(secondary)*	**Scale, Proportion, and Quantity** • The significance of a phenomenon is dependent on the scale, proportion, and quantity at which it occurs.
Clarifications and Content Limits	**Clarification Statements** • Emphasis is on the energy transfer mechanisms that allow energy from nuclear fusion in the sun's core to reach Earth. • Examples of evidence for the model include observations of the masses and lifetimes of other stars, as well as the ways that the sun's radiation varies due to sudden solar flares ("space weather"), the 11-year sunspot cycle, and non-cyclic variations over centuries. **Content Limits** • Assessment does not include details of the atomic and sub-atomic processes involved with the sun's nuclear fusion.		
Science Vocabulary Students are Expected to Know	sunspot cycle, solar maximum, solar minimum, sunspots, solar flares, UV radiation, IR radiation, convection, nuclear fusion, core, atmosphere, solar storm, luminosity		
Science Vocabulary Students are Not Expected to Know	photosphere, chromosphere, corona, coronal mass ejections		
Phenomena			
Context/ Phenomena	Some example phenomena for HS-ESS1-1: • The habitable zone in our solar system currently contains both Earth and Mars. In the future it will contain a different set of planets. • The sun's current surface temperature is about 5,800 K. In 5 billion years, the sun's surface temperature will cool to 3,500 K. • The sun is 40% brighter, 6% larger than 5% hotter than it was 5 billion years ago. • The Earth's atmosphere will contain more water vapor and the oceans will contain less water in a few billion years.		
This Performance Expectation and associated Evidence Statements support the following Task Demands.			
Task Demands			
1. Organize and/or arrange (e.g., using illustrations and/or labels), summarize or make inferences about data to highlight trends, patterns, or correlations.			

2.	Identify patterns or evidence in the data that supports inferences about the lifespan of the sun or the transfer of energy from the sun to the earth.
3.	Select or identify from a collection of potential model components, including distractors, the components needed for a model that illustrates the lifespan of the sun or the transfer of energy from the sun to the earth.
4.	Construct or complete a model capable of illustrating the lifespan of the sun or the transfer of energy from the sun to the earth.
5.	Manipulate the components of a model to demonstrate the changes, properties, processes, and/or events that are relevant to the lifespan of the sun or the transfer of energy from the sun to the earth.
6.	Identify missing components, relationships, or other limitations of the model.
7.	Make predictions about the effects of changes in the sun or in the transfer of energy from the sun to the earth. Predictions can be made by manipulating model components, completing illustrations, or selecting from lists with distractors.

<table>
<tr><td>Performance Expectation</td><td colspan="3">HS-ESS1-2
Construct an explanation of the Big Bang theory based on astronomical evidence of light spectra, motion of distant galaxies, and composition of matter in the universe.</td></tr>
<tr><td>Dimensions</td><td>Constructing Explanations and Designing Solutions
• Construct an explanation based on valid and reliable evidence obtained from a variety of sources (including students' own investigations, theories, simulations, peer review) and the assumption that theories and laws that describe the natural world operate today as they did in the past and will continue to do so in the future.</td><td>ESS1.A: The Universe and Its Stars
• The study of stars' light spectra and brightness is used to identify compositional elements of stars, their movements, and their distances from Earth.
• The Big Bang theory is supported by observations of distant galaxies receding from our own, of the measured composition of stars and non-stellar gases, and of the maps of spectra of the primordial radiation (cosmic microwave background) that still fills the universe.
• Other than the hydrogen and helium formed at the time of the Big Bang, nuclear fusion within stars produces all atomic nuclei lighter than and including iron, and the process releases electromagnetic energy. Heavier elements are produced when certain massive stars achieve a supernova stage and explode.

PS4.B: Electromagnetic Radiation
• Atoms of each element emit and absorb characteristic frequencies of light. These characteristics allow identification of the presence of an element, even in microscopic quantities. (secondary)</td><td>Energy and Matter
• Energy cannot be created or destroyed–only moved between one place and another place, between objects and/or fields, or between systems.</td></tr>
<tr><td>Clarifications and Content Limits</td><td colspan="3">Clarification Statements
• Emphasis is on the astronomical evidence of the redshift of light from galaxies as an indication that the universe is currently expanding, the cosmic microwave background as the remnant radiation from the Big Bang, and the observed composition of ordinary matter of the universe, primarily found in stars and interstellar gases (from the spectra of electromagnetic radiation from stars), which matches that predicted by the Big Bang theory (3/4 hydrogen and 1/4 helium).</td></tr>
<tr><td>Science Vocabulary Students are Expected to Know</td><td colspan="3">Recessional velocity, galaxy, star, galaxy cluster, spectrum, spectra, wavelength, frequency, Doppler Effect, redshift, blueshift, light years, big bang theory, helium, emission, absorption</td></tr>
<tr><td>Science Vocabulary Students are Not Expected to Know</td><td colspan="3">Cosmological redshift, Hubble Law, photometric redshift, spectroscopy</td></tr>
<tr><td colspan="4">Phenomena</td></tr>
<tr><td>Context/ Phenomena</td><td colspan="3">Some example Phenomena for HS-ESS1-2:
• The farthest known galaxy has a greater recessional velocity than the farthest known quasar.</td></tr>
</table>

	• The spectrum of NGC450 shows a greater abundance of elements heavier than helium than does the spectrum of NGC60 • Two galaxy clusters observed in opposite parts of the sky both contain galaxies with about the same chemical composition: 75% hydrogen and 25% helium. • A galaxy in the constellation Cetus is moving away from us at a different speed than another galaxy in the adjacent constellation Pisces.

This Performance Expectation and associated Evidence Statements support the following Task Demands.
Task Demands
1. Articulate, describe, illustrate, or select the relationships, interactions, and/or processes to be explained. This may entail organizing, interpreting and analyzing data, making calculations, and sorting relevant from irrelevant information or features.
2. Identify evidence that supports and/or does not support the Big Bang Theory.
3. Describe, select, or identify components of the Big Bang Theory supported by given evidence.
4. Use an explanation of the Big Bang Theory to predict how the universe will continue to change over time.
5. Construct an explanation based on evidence that explains how particular aspects of the Big Bang Theory are supported by empirical observations of the universe.
6. Identify and justify additional pieces of evidence that would help distinguish among competing hypotheses.

<table>
<tr><td>Performance Expectation</td><td colspan="3">HS-ESS1-3
Communicate scientific ideas about the way stars, over their life cycle, produce elements.</td></tr>
<tr><td>Dimensions</td><td>Obtaining, Evaluating, and Communicating Information
• Communicate scientific ideas (e.g. about phenomena and/or the process of development and the design and performance of a proposed process or system) in multiple formats (including orally, graphically, textually, and mathematically).</td><td>ESS1.A: The Universe and Its Stars
• The study of stars' light spectra and brightness is used to identify compositional elements of stars, their movements, and their distances from Earth.
• Other than the hydrogen and helium formed at the time of the Big Bang, nuclear fusion within stars produces all atomic nuclei lighter than and including iron, and the process releases electromagnetic energy. Heavier elements are produced when certain massive stars achieve a supernova stage and explode.</td><td>Energy and Matter
• In nuclear processes, atoms are not conserved, but the total number of protons plus neutrons is conserved.</td></tr>
<tr><td>Clarifications and Content Limits</td><td colspan="3">Clarification Statements
• Emphasis is on the way nucleosynthesis, and therefore the different elements created, varies as a function of the mass of a star and the stage of its lifetime.

Content Limits
• Details of the many different nucleosynthesis pathways for stars of differing masses are not assessed.
• Include basic/simplified nucleosynthesis reactions:
  • Hydrogen fuses into helium
  • Helium fuses into carbon
  • Carbon fuses into oxygen
  • Oxygen fuses into silicon
  • Silicon fuses into iron
• Exclude complex nucleosynthesis reactions and details:
  ○ CNO cycle
  ○ Neutron-capture (r-process and s-process)
  ○ Proton-capture: Rp-process
  ○ Photo-disintegration: P-process
  ○ Other details about radiation or particles – focus on conservation of nucleons</td></tr>
<tr><td>Science Vocabulary Students are Expected to Know</td><td colspan="3">main sequence, nucleosynthesis, nuclear reactions, fission, fusion, nucleons, proton, neutron, , , gamma rays, neutrinos, red giant, blue giant, white dwarf, planetary nebular, supernova, supernova remnant, globular cluster, open , exothermic reactions, endothermic reactions, emissions spectrum, absorption spectrum, emission lines, absorption lines, H-R Diagram</td></tr>
<tr><td>Science Vocabulary Students are Not Expected to Know</td><td colspan="3">Neutron-capture, proton-capture, photo-disintegration, CNO cycle, radiogenesis</td></tr>
<tr><td colspan="4">Phenomena</td></tr>
<tr><td>Context/ Phenomena</td><td colspan="3">Some example phenomenon for HS-ESS1-3:</td></tr>
</table>

	• Two larger stars, Spica and Pollux are eight times larger than the sun. However, Spica is 420 times brighter and 6 times more massive than Pollux. • Procyon is a 1.5 solar mass star and is 8 times brighter than the sun. Aldebaran is a star of similar mass but Aldebaran is 425 times brighter than the sun. • The stars in a globular cluster (old low mass stars) are red and show few absorption lines in their spectra while the stars in an open cluster (young high mass stars) are blue and show many absorption lines in their spectra. • In the core of some stars, carbon can fuse into neon, sodium or magnesium.

This Performance Expectation and associated Evidence Statements support the following Task Demands.
Task Demands
1. Illustrate, model or make calculations involving the nucleosynthesis process in stars of different mass, different luminosity, different age or different evolutionary stage using graphs, diagrams, text and mathematical models.
2. Compare and contrast the nucleosynthesis processes of stars of different mass, different luminosity, different age or different evolutionary stage using graphs, diagrams, text and mathematical models.
3. Make predictions about nucleosynthesis processes given changes or differences in other stellar characteristics.
4. Identify and communicate evidence supporting an explanation regarding the relationship between stellar properties and age, in particular how those stellar properties change over time.
5. Synthesize an explanation regarding the relationship between stellar properties and age, in particular how those stellar properties change over time.

Performance Expectation	**HS-ESS1-4** Use mathematical or computational representations to predict the motion of orbiting objects in the solar system.		
Dimensions	**Using Mathematical and Computational Thinking** • Use mathematical or computational representations of phenomena to describe explanations.	**ESS1.B: Earth and the Solar System** • Kepler's laws describe common features of the motions of orbiting objects, including their elliptical paths around the sun. Orbits may change due to the gravitational effects from, or collisions with, other objects in the solar system.	**Scale, Proportion, and Quantity** • Algebraic thinking is used to examine scientific data and predict the effect of a change in one variable on another (e.g., linear growth vs. exponential growth).
Clarifications and Content Limits	**Clarification Statements** • Emphasis is on Newtonian gravitational laws governing orbital motions, which apply to human-made satellites as well as planets, moons, rings, asteroids, and comets. • The term "satellite" can be used to describe both man-made and natural objects that orbit another object. **Content Limits** • Mathematical representations for the gravitational attraction of bodies and Kepler's Laws of orbital motions should not deal with systems of more than two bodies, nor involve calculus. • Comparing different orbiting bodies is acceptable as long as each system only contains two bodies (example: satellite 1 orbiting Earth compared to satellite 2 orbiting Earth). • Students will be given the Law of Gravitation to make calculations but should know/apply Kepler's laws conceptually. These laws are: 1. Orbits are elliptical; 2. Line connecting orbiting body and parent body sweeps out equal areas in equal time; 3. (Orbital period)^2 is proportional to (semi-major axis distance)^3.		
Science Vocabulary Students are Expected to Know	Gravitation, orbit, revolution, rotation, period, semi-major axis, eccentricity, semi-minor axis, focus, foci, ellipse, gravitational constant, astronomical unit, satellite		
Science Vocabulary Students are Not Expected to Know	Aphelion, perihelion, angular momentum		
Phenomena			
Context/ Phenomena	Some sample phenomena for HS-ESS1-4: • The International Space Station orbits Earth at an altitude of 250 miles with a speed of 5 miles per second while a global positioning system satellite orbits ten times as far and half as fast. • China's Tiangong space station's orbital speed can no longer be controlled. It is expected to burn up in the atmosphere as it falls to the Earth. • The shape of Comet Shoemaker-Levy 9's orbit changed just before it collided with Jupiter in 1994. • In 100 years, the moon will be about half a meter further from Earth and Earth's rotation will be 2 miliseconds slower.		

This Performance Expectation and associated Evidence Statements support the following Task Demands.	
Task Demands	
1. Make simple calculations using given data to calculate or estimate the motion of orbiting objects (satellites).	
2. Illustrate, graph, or identify relevant features or data that can be used to calculate, estimate or make inferences about the motion of satellites.	
3. Calculate or estimate properties of motions for a satellite and the object it orbits based on data from one or more sources.	
4. Select or construct relationships between a satellite and the object it orbits based on data from one or more sources.	
5. Compile, from given information, the particular data needed for a particular inference about the motion of a satellite. This can include sorting out the relevant data from the given information.	
6. Construct or identify an inference that can be made based on data from one or more sources.	

<table>
<tr><td>Performance Expectation</td><td colspan="3">HS-ESS1-5
Evaluate evidence of the past and current movements of continental and oceanic crust and the theory of plate tectonics to explain the ages of crustal rocks.</td></tr>
<tr><td>Dimensions</td><td>Engaging in Argument from Evidence
• Evaluate evidence behind currently accepted explanations or solutions to determine the merits of arguments.</td><td>ESS1.C: The History of Planet Earth
• Continental rocks, which can be older than 4 billion years, are generally much older than the rocks of the ocean floor, which are less than 200 million years old.

ESS2.B: Plate Tectonics and Large-Scale System Interactions
• Plate tectonics is the unifying theory that explains the past and current movements of the rocks at Earth’s surface and provides a framework for understanding its geologic history. (secondary)

PS1.C: Nuclear Processes
• Spontaneous radioactive decays follow a characteristic exponential decay law. Nuclear lifetimes allow radiometric dating to be used to determine the ages of rocks and other materials. (secondary)</td><td>Patterns
• Empirical evidence is needed to identify patterns.</td></tr>
<tr><td>Clarifications and Content Limits</td><td colspan="3">Clarification Statements
• Emphasis is on the ability of plate tectonics to explain the ages of crustal rocks.
• Examples include evidence of the ages of oceanic crust increasing with distance from mid-ocean ridges (a result of plate spreading) and the ages of North American continental crust decreasing with distance away from a central ancient core of the continental plate (a result of past plate interactions).

Content Limits
• Students do not need to calculate radioactive decay rates.
• <u>Students do not need to know:</u> names of supercontinents, names of fault lines, names of tectonic plates</td></tr>
<tr><td>Science Vocabulary Students are Expected to Know</td><td colspan="3">Convergence, divergence, sedimentary, metamorphic, igneous, volcanic, crust, mantle, core, mid ocean ridge, trench</td></tr>
<tr><td>Science Vocabulary Students are Not Expected to Know</td><td colspan="3">Isotope, anticline, syntacline</td></tr>
<tr><td colspan="4">Phenomena</td></tr>
<tr><td>Context/ Phenomena</td><td colspan="3">Some example phenomena for HS-ESS1-5:
• Rocks near Bildudalur Iceland were formed about about 16 million years ago, rocks near Geysir Iceland were formed about 3.3 million years ago.
• The patterns of magnetic reversals on the youngest continental rock columns are the same as the pattern of magnetic reversals found at the center of the Mid-Atlantic ridge.
• Iceland gains about 1.8 centimeters of land surface per year.
• From 1996 to 2016, Mount St. Elias has gotten 0.08 meters taller.</td></tr>
</table>

This Performance Expectation and associated Evidence Statements support the following Task Demands.	
Task Demands	
1. Based on the provided data or information, identify the explanation that could explain the age difference in continental and oceanic crust.	
2. Identify and/or explain the claims, evidence, and reasoning supporting the explanation that tectonic plates have moved over time.	
3. Identify and/or describe additional relevant evidence not provided that would support or clarify the explanation of the movement of tectonic plates and/or the ages of rocks.	
4. Evaluate the strengths and weaknesses of a claim to explain the theory of plate tectonics and the ages of rocks.	
5. Analyze and/or interpret evidence and its ability to support the explanation that plate tectonics or radioactive decay can determine the age of a rock.	
6. Provide and/or evaluate reasoning to support the explanation that volcanoes, mountains and earthquakes are formed/caused as a result of plate tectonics	

<table>
<tr><td>Performance Expectation</td><td colspan="3">HS-ESS1-6
Apply scientific reasoning and evidence from ancient Earth materials, meteorites, and other planetary surfaces to construct an account of Earth’s formation and early history.</td></tr>
<tr><td>Dimensions</td><td>Constructing Explanations and Designing Solutions
• Apply scientific reasoning to link evidence to the claims to assess the extent to which the reasoning and data support the explanation or conclusion.</td><td>ESS1.C: The History of Planet Earth
• Although active geologic processes, such as plate tectonics and erosion, have destroyed or altered most of the very early rock record on Earth, other objects in the solar system, such as lunar rocks, asteroids, and meteorites, have changed little over billions of years. Studying these objects can provide information about Earth’s formation and early history.

PS1.C: Nuclear Processes
• Spontaneous radioactive decays follow a characteristic exponential decay law. Nuclear lifetimes allow radiometric dating to be used to determine the ages of rocks and other materials. (secondary)</td><td>Stability and Change
• Much of science deals with constructing explanations of how things change and how they remain stable.</td></tr>
<tr><td>Clarifications and Content Limits</td><td colspan="3">Clarification Statements
• Emphasis is on using available evidence within the solar system to reconstruct the early history of Earth, which formed along with the rest of the solar system 4.6 billion years ago.
• Examples of evidence include the absolute ages of ancient materials (obtained by radiometric dating of meteorites, moon rocks, and Earth’s oldest minerals), the sizes and compositions of solar system objects, and the impact cratering record of planetary surfaces.</td></tr>
<tr><td>Science Vocabulary Students are Expected to Know</td><td colspan="3">Plate tectonics, radiometric dating, isotope, continental crust, oceanic crust, lithosphere, asthenosphere, cycle, bedrock, ocean trench, sedimentation, convection current, ancient core, inner core, mantle, nuclear, ocean ridge, sea-floor spreading</td></tr>
<tr><td>Science Vocabulary Students are Not Expected to Know</td><td colspan="3">Nebular hypothesis, planetesimals, solar nebula, bolide impacts,</td></tr>
<tr><td colspan="4">Phenomena</td></tr>
<tr><td>Context/ Phenomena</td><td colspan="3">Some sample phenomena for HS-ESS1-6:
• A thin section of a rock from western Australia is examined under a microscope and elongate crystals are observed.
• A rock from Earth and a rock from Mars are the same age.
• When astronauts returned to Earth with rocks from the moon, they were all very old. A rock found in the Great Lakes Region of North America is very old, but rock found in Iceland are all relatively young. Meteor Crater is a large depression, with a depth of 170m, in an otherwise flat area of Arizona.</td></tr>
<tr><td colspan="4">This Performance Expectation and associated Evidence Statements support the following Task Demands.</td></tr>
<tr><td colspan="4">Task Demands</td></tr>
<tr><td colspan="4">1. Articulate, describe, illustrate, or select the relationships, interactions, and/or processes to be explained. This may entail sorting relevant from irrelevant information or features.</td></tr>
</table>

2. Express or complete a causal chain explaining Earth's formation and/or early history. This may include indicating directions of causality in an incomplete model such as a flow chart or diagram, or completing cause and effect chains.*
3. Identify evidence supporting the inference of causation that is expressed in a causal chain.
4. Describe, identify, and/or select information needed to support an explanation about the formation of Earth and its early history.
5. Construct an explanation based on evidence and scientific reasoning that explains the formation of Earth and its early history. *****

*denotes those task demands which are deemed appropriate for use in stand-alone item development

<table>
<tr><td>Performance Expectation</td><td colspan="3">HS-ESS2-1
Develop a model to illustrate how Earth’s internal and surface processes operate at different spatial and temporal scales to form continental and ocean-floor features.</td></tr>
<tr><td>Dimensions</td><td>Developing and Using Models
• Develop a model based on evidence to illustrate the relationships between systems or between components of a system.</td><td>ESS2.A. Earth Materials and Systems
• Earth’s systems, being dynamic and interacting, cause feedback effects that can increase or decrease the original changes.

ESS2.B. Plate Tectonics and Large-Scale System Interactions
• Plate tectonics is the unifying theory that explains the past and current movements of the rocks at Earth’s surface and provides a framework for understanding its geologic history.
• Plate movements are responsible for most of continental and ocean-floor features and for the distribution of most rocks and minerals within Earth’s crust.</td><td>Stability and Change
• Change and rates of change can be quantified and modeled over very short or very long periods of time. Some system changes are irreversible.</td></tr>
<tr><td>Clarifications and Content Limits</td><td colspan="3">Clarification Statements
• Emphasis is on how the appearance of land features (such as mountains, valleys, and plateaus) and sea floor features (such as trenches, ridges, and seamounts) are a result of both constructive forces (such as volcanism, tectonic uplift, and orogeny) and destructive mechanisms (such as weathering, mass wasting, and coastal erosion).

Content Limits
• <u>Students do not need to know:</u> the details of the formation of specific geographic features of Earth’s surface.</td></tr>
<tr><td>Science Vocabulary Students are Expected to Know</td><td colspan="3">Tectonic uplift, seismic waves, feedback effect, irreversible, Earth’s magnetic field, electromagnetic radiation, inner core, outer core, mantle, continental crust, oceanic crust, sea-floor spreading, isotope, thermal convection, radioactive decay, rock composition, continental boundary, ocean trench, recrystallization, nuclear, geochemical reaction, mass wasting</td></tr>
<tr><td>Science Vocabulary Students are Not Expected to Know</td><td colspan="3">Geomorphology, anticline, syncline, monocline</td></tr>
<tr><td colspan="4">Phenomena</td></tr>
<tr><td>Context/ Phenomena</td><td colspan="3">Some sample phenomena for HS-ESS2-1:
• A limestone cliff that contains Cambrian-aged fossils extends several hundred feet above the surface of the ocean. A large section of the cliff has collapsed.
• An oceanic trench 10,000 is meters below sea level. Inland, 200km away, a chain of active volcanoes is present.
• 1.8 billion year old rocks in the Black Hills of South Dakota are capped by 10,000 year old gravel terraces.
• A photograph from March shows large Precambrian-aged pink granite boulder at the top of a 100 m tall hill. A photograph in April shows the same boulder sitting in a pile of soil and sediment in the valley below the hill.</td></tr>
</table>

This Performance Expectation and associated Evidence Statements support the following Task Demands.
Task Demands
1. Select or identify from a collection of potential model components, including distractors, the components that are relevant for explaining the phenomenon. Components might include different rock types, rates of uplift and erosion, surface environments on Earth where these processes occur and where different rock types exist, and layers within Earth where these processes occur. Sources of energy (radiation, convection) that drive the cycling (but *not* the creation of) matter should also be included as components.*
2. Manipulate the components of a model to demonstrate the changes, properties, processes, and/or events that act to result in the phenomenon of Earth's internal and surface processes.
3. Make predictions about the effects of changes in the magnitude and/or rate of Earth's internal and surface properties. Predictions can be made by manipulating model components, completing illustrations, or selecting from lists with distractors.
4. Given models or diagrams of land features, internal and surface processes, identify factors that affect constructive and destructive forces, feedback effects and how they vary in different scenarios OR identify the constructive and destructive mechanisms that operate at different spatial and temporal time scales and how this causes changes in the appearance of continental and ocean-floor features.
5. Identify missing components, relationships, or other limitations of the model of how Earth's internal and surface processes form continental and ocean-floor features.
6. Describe, identify, or select the relationships among components of a model that describe the formation of continental and ocean-floor features with respect to spatial and temporal variability in internal and external surface processes or explains how changes in these processes affect the formation of continental and ocean-floor features.*
7. Express or complete a causal chain explaining how changes in the flow of energy (interval vs. surface processes) affect the formation of continental and ocean-floor features. This may include indicating directions of causality in an incomplete model such as a flow chart or diagram, or completing cause and effect chains.

*denotes those task demands which are deemed appropriate for use in stand-alone item development

<table>
<tr><td>Performance Expectation</td><td colspan="3">HS-ESS2-2
Analyze geoscience data to make the claim that one change to Earth’s surface can create feedbacks that cause changes to Earth’s systems.</td></tr>
<tr><td>Dimensions</td><td>Analyzing and Interpreting Data
• Analyze data using tools, technologies and/or models (e.g. computational, mathematical) in order to make valid and reliable scientific claims or determine an optimal design.</td><td>ESS2.A: Earth Materials and Systems
• Earth’s systems, being dynamic and interacting, cause feedback effects that can increase or decrease the original changes.

ESS2.D: Weather and Climate
• The foundation for the Earth’s global climate systems is the electromagnetic radiation from the sun, as well as its reflection, absorption, storage and redistribution among the atmosphere, ocean, and land systems, and this energy’s re-radiation into space.</td><td>Stability and Change
• Feedback (negative or positive) can stabilize or destabilize a system.</td></tr>
<tr><td>Clarifications and Content Limits</td><td colspan="3">Clarification Statements
• Examples should include climate feedbacks, such as:
  • An increase in greenhouse gases causes a rise in global temperatures that melts glacial ice, which reduces the amount of sunlight reflected from Earth’s surface, increasing surface temperatures and further reducing the amount of ice.
  • Loss of ground vegetation causes an increase in water runoff and soil erosion
  • Damned rivers increase groundwater recharge, decrease sediment transport, and increase coastal erosion
  • Loss of wetlands causes a decrease in local humidity that further reduces the wetland extent.

Content Limits
• <u>Students do not need to know:</u>
  o Specifically which gases are greenhouse gases.
  o Composition of the atmosphere</td></tr>
<tr><td>Science Vocabulary Students are Expected to Know</td><td colspan="3">Ocean circulation, biosphere, feedback effect, atmospheric circulation, convection cycle, greenhouse gas, geoscience, sea level, mean surface temperature, methane</td></tr>
<tr><td>Science Vocabulary Students are Not Expected to Know</td><td colspan="3">Electromagnetic radiation, probabilistic, irreversible, geoengineering, ozone, pollutant, acidification</td></tr>
<tr><td colspan="4">Phenomena</td></tr>
<tr><td>Context/ Phenomena</td><td colspan="3">Some example phenomena for HS-ESS2-2:
• Farming causes the loss of forest in the Amazon. This leads to an increase in erosion and water runoff, which leads to more forest loss.
• Loss of wetlands causes a decrease in local humidity that further reduces the wetland extent.
• As the Permafrost in the Artic melts, methane is released into the atmosphere. Methane, a greenhouse gas, traps heat causing the Earth to heat up, leading to more Permafrost melting.</td></tr>
</table>

	• Increased CO_2 in the atmosphere warms the oceans. Warmer oceans take up less CO_2 than cooler oceans, further increasing atmospheric temperature.
This Performance Expectation and associated Evidence Statements support the following Task Demands.	
Task Demands	
1. Organize and/or arrange (e.g., using illustrations and/or labels), or summarize data to highlight trends, patterns, or correlations in how changes to Earth's surface can create feedbacks that affect Earth's systems.	
2. Generate/construct graphs, tables, or assemblages of illustrations and/or labels of data that document patterns, trends, or correlations in how changes to Earth's surface can create feedbacks that affect Earth's systems. This may include sorting out distractors.	
3. Use relationships identified in the data to predict how changing the Earth's surfaces affects the feedback loop.	
4. Identify patterns or evidence in the data that supports inferences about how the altering of Earth's surface will affect the Earth in the long term.	

Performance Expectation	**HS-ESS2-3** Develop a model based on evidence of Earth's interior to describe the cycling of matter by thermal convection.		
Dimensions	**Develop and Using Models** • Develop a model based on evidence to illustrate the relationships between systems or between components of a system.	**ESS2.A: Earth Materials and Systems** • Evidence from deep probes and seismic waves, reconstructions of historical changes in Earth's surface and its magnetic field, and an understanding of physical and chemical processes lead to a model of Earth with a hot but solid inner core, a liquid outer core, a solid mantle and crust. Motions of the mantle and its plates occur primarily through thermal convection, which involves the cycling of matter due to the outward flow of energy from Earth's interior and gravitational movement of denser materials toward the interior. **ESS2.B: Plate Tectonics and Large-Scale System Interactions** • The radioactive decay of unstable isotopes continually generates new energy within Earth's crust and mantle, providing the primary source of the heat that drives mantle convection. Plate tectonics can be viewed as the surface expression of mantle convection. **PS4.A: Wave Properties** • Geologists use seismic waves and their reflection at interfaces between layers to probe structures deep in the planet *(secondary)*	**Energy and Matter** • Energy drives the cycling of matter within and between systems.
Clarifications and Content Limits	**Clarification Statements** • Emphasis is on both a one-dimensional model of Earth, with radial layers determined by density, and a three-dimensional model, which is controlled by mantle convection and the resulting plate tectonics. • Examples of evidence include maps of Earth's three-dimensional structure obtained from seismic waves, records of the rate of change of Earth's magnetic field (as constraints on convection in the outer core), and identification of the composition of Earth's layers from high-pressure laboratory experiments.		
Science Vocabulary Students are Expected to Know	Convection, radioactive, inner core, outer core, isotope, mantle, seismic wave, Geochemical reaction, geoscience, molten rock, Earth's elements, Earth's internal energy sources, geochemical cycle, tectonic uplift		
Science Vocabulary Students are Not Expected to Know	Geoneutrino, primodial heat		
Phenomena			
Context/ Phenomena	Some example phenomena for HS-ESS2-3: • The temperature of the water in a hot spring in Iceland is around 100°F. The average air temperature in Iceland is about 52°F.		

	• The average heat flow from the Earth's interior is 80 mWm^{-2}. The heat flow of a volcano on Hawaii is ~400 mWm^{-2}. • The total heat transfer from the Earth to space is 44 terawatts. Radioactive decay of unstable isotopes contributes 20 terawatts from Earth's interior. (KamLAND Collaboration, 2011). • In the central valley of California, the temperature at 5 meters below the ground is 2°C warmer than the temperature at the surface. In northern Oregon near Mt. Hood, the temperature 5 meters underground is 10°C warmer than the temperature at the surface.

This Performance Expectation and associated Evidence Statements support the following Task Demands.
Task Demands
1. Select or identify from a collection of potential model components, including distractors, the components needed to model the phenomenon. Components might include the structure of the Earth, the cycling of matter and/or energy, or instruments used to measure seismic waves.
2. Assemble or complete, from a collection of potential model components, an illustration or flow chart that is capable of representing the structure and the flow of matter/energy from the Earth's interior. This does not include labeling an existing diagram.
3. Manipulate the components of a model to demonstrate the changes, properties, processes, and/or events that act to result in the phenomenon.
4. Make predictions about the effects of changes in the cycling of matter and energy. Predictions can be made by manipulating model components, completing illustrations, or selecting from lists with distractors.
5. Given models or diagrams of the earth's interior, identify the chemical and physical properties of the Earth's structure that cause the cycling of matter.
6. Identify missing components, relationships, or other limitations of the model.
7. Describe, select, or identify the relationships among components of a model that describe the cycling of matter within Earth's interior.

<table>
<tr><td>Performance Expectation</td><td colspan="3">HS-ESS2-4
Use a model to describe how variations in the flow of energy into and out of Earth’s systems result in changes in climate.</td></tr>
<tr><td>Dimensions</td><td>Developing and Using Models
• Use a model to provide mechanistic accounts of phenomena.</td><td>ESS1.B: Earth and the Solar System
• Cyclical changes in the shape of Earth’s orbit around the sun, together with changes in the tilt of the planet’s axis of rotation, both occurring over hundreds of thousands of years, have altered the intensity and distribution of sunlight falling on Earth. These phenomena cause a cycle of ice ages and other gradual climate changes. (secondary)

ESS2.A: Earth Materials and Systems
• The geologic record shows that changes to global and regional climate can be caused by interactions among changes in the sun’s energy output of Earth’s orbit, tectonic events, ocean circulation, volcanic activity, glaciers, vegetation, and human activities. These changes can occur on a variety of timescales from sudden (e.g., volcanic ash clouds) to intermediate (ice ages) to very long-term tectonic cycles.

ESS2.D: Weather and Climate
• The foundation for Earth’s global climate system is the electromagnetic radiation from the sun, as well as its reflection, absorption, storage, and redistribution among the atmosphere, ocean, and land systems and this energy’s re-radiation into space.</td><td>Cause and Effect
• Empirical evidence is required to differentiate between cause and correlation and make claims about specific causes and effects.</td></tr>
<tr><td>Clarifications and Content Limits</td><td colspan="3">Clarification Statements
• Examples of the causes of climate change differ by time scale, over 1-10 years: large volcanic eruption, ocean circulation; 10-100s of years: changes in human activity, ocean circulation, solar output; 10-100s of thousands of years: changes to Earth's orbit and the orientation of its axis; and 10-100s of millions of years: long-term changes in atmospheric composition.

Content Limits
• Assessment of the results of changes in climate is limited to changes in surface temperatures, precipitation patterns, glacial ice volumes, sea levels, and biosphere distribution.
• <u>Students do not need to know:</u> chemical mechanisms of fossil fuel combustion or ozone depletion</td></tr>
<tr><td>Science Vocabulary Students are Expected to Know</td><td colspan="3">Interdependence, solar radiation, solar flare, biosphere, atmospheric circulation, ocean circulation, climatic patternsea level, glacier, atmospheric composition, hydrosphere, greenhouse gas, fossil fuel, combustion</td></tr>
<tr><td>Science Vocabulary Students are</td><td colspan="3">Acidification, cryosphere</td></tr>
</table>

Not Expected to Know	
Phenomena	
Context/ Phenomena	Some example phenomena for HS-ESS2-4: • Temperatures were warmer in 1990 than in the 5 previous years. In 1992 and 1993, the global temperatures were 1°F cooler than in 1991. (volcanic eruption of Mount Pinatubo) • 11,000 years ago large portions of the northern United States contained glaciers. Today, very little of this area contains glaciers. (changes to Earth's orbit) • Earth experiences 4 distinct seasons. Venus does not experience distinct seasons. (tilt of planet's axis) • 25,000 years ago, the level of carbon dioxide in the atmosphere was around 180 parts per million (ppm). Today, carbon dioxide levels exceed 400 ppm. (atmospheric composition)
This Performance Expectation and associated Evidence Statements support the following Task Demands.	
Task Demands	
1. Select or identify from a collection of potential model components, including distractors, the components that are relevant for explaining the phenomenon. Components might include factors that affect the input, storage, redistribution, and output of energy in Earth's systems.	
2. Manipulate the components of a model to demonstrate the changes, properties, processes, and/or events that act to result in the phenomenon of the flow of energy in Earth's systems.	
3. Make predictions about the effects of changes in energy flow on Earth's climate.	
4. Given models or diagrams of energy flow in Earth's systems, identify factors that affect energy input, output, storage, and redistribution and how they change in different scenarios OR identify the changes in energy flow that cause changes in Earth's climate.	
5. Identify missing components, relationships, or other limitations of the model of energy flow in Earth's systems.	
6. Describe, identify, or select the relationships among components of a model that describe changes in the flow of energy in Earth's systems or explains how changes in energy flow affect climate.	
7. Express or complete a causal chain explaining how changes in the flow of energy in Earth's systems affects climate. This may include indicating directions of causality in an incomplete model such as a flow chart or diagram, or completing cause and effect chains.	

<table>
<tr><td>Performance Expectation</td><td colspan="3">HS-ESS2-5
Plan and conduct an investigation of the properties of water and its effects on Earth materials and surface processes.</td></tr>
<tr><td>Dimensions</td><td>Planning and Carrying Out Investigations
• Plan and conduct an investigation individually and collaboratively to produce data to serve as the basis for evidence, and in the design decide on types, how much, and accuracy of data needed to produce reliable measurements and consider limitations on the precision of the data (e.g., the number of trials, cost, risk, time), and refine the design accordingly.</td><td>ESS2.C: The Roles of Water in Earth's Surface Processes
• The abundance of liquid water on Earth's surface and its unique combination of physical and chemical properties are central to the planet's dynamics. These properties include water's exceptional capacity to absorb, store, and release large amounts of energy, transmit sunlight, expand upon freezing, dissolve and transport materials, and lower the viscosities and melting points of rocks.</td><td>Structure and Function
• The functions and properties of natural and designed objects and systems can be inferred from their overall structure, the way their components are shaped and used, and the molecular substructures of its various materials.</td></tr>
<tr><td>Clarifications and Content Limits</td><td colspan="3">Clarification Statements
• Emphasis is on mechanical and chemical investigations with water and a variety of solid materials to provide evidence for the connections between the hydrologic cycle and system interactions commonly known as the rock cycle.
• Examples of mechanical investigations include stream transportation and deposition using a stream table, erosion using variations in soil moisture content, and frost wedging by the expansion of water as it freezes.
• Examples of chemical investigations include chemical weathering and recrystallization (by testing the solubility of different materials) or melt generation (by examining how water lowers the melting temperature of most solids).

Content Limits
• The abundance of liquid water on Earth's surface and its unique combination of physical and chemical properties are central to the planet's dynamics. These properties include water's exceptional capacity to absorb, store, and release large amounts of energy; transmit sunlight; expand upon freezing; dissolve and transport materials; and lower the viscosities and melting points of rocks.</td></tr>
<tr><td>Science Vocabulary Students are Expected to Know</td><td colspan="3">Viscosity, melting point, freezing point, absorption, dissolve, hydrologic cycle, rock cycle, stream transportation, stream deposition, stream table, erosion, soil moisture content, frost wedging, chemical weathering, solubility, mechanical erosion, heat capacity, density, molecular structure, sediment, cohesion, polarity.</td></tr>
<tr><td>Science Vocabulary Students are Not Expected to Know</td><td colspan="3"></td></tr>
<tr><td colspan="4">Phenomena</td></tr>
</table>

Context/ Phenomena	Some example phenomena for HS-ESS2-5: • In a cave in Guam, sections of stalactites that formed during seasons of high rainfall contain a lower ratio of the isotopes oxygen-18 to oxygen-16 than sections of the stalactites that formed during seasons of low rainfall. • Wookey Hole Caves have about 4,000 meters of cave system in a rock formation. • The Colorado River runs through the rock formation known as the Grand Canyon.

This Performance Expectation and associated Evidence Statements support the following Task Demands.
Task Demands
1. Identify from a list, including distractors, the materials/tools needed for an investigation of the properties of water and its effects on Earth’s materials and surface processes.
2. Identify the outcome data that should be collected in an investigation of the properties of water and its effects on Earth’s materials and surface processes.
3. Evaluate the sufficiency and limitations of data collected to explain the effects of water on Earth’s materials and surface processes.
4. Make and/or record observations about the chemical and/or physical properties of liquid water and its effects on Earth’s materials.
5. Interpret and/or communicate the data from an investigation of the effect of water on Earth’s materials and surface processes.
6. Explain or describe the causal processes that lead to the observed effects of water.
7. Select, describe, or illustrate a prediction made by applying the findings from an investigation of the effects of water on Earth’s materials and surface processes.

Performance Expectation	**HS-ESS2-6** Develop a quantitative model to describe the cycling of carbon among the hydrosphere, atmosphere, geosphere, and biosphere.		
Dimensions	**Developing and Using Models** • Develop a model based on evidence to illustrate the relationships between systems or between components of a system.	**ESS2.D Weather and Climate** • Gradual atmospheric changes were due to plants and other organisms that captured carbon dioxide and released oxygen. • Changes in the atmosphere due to human activity have increased carbon dioxide concentrations and thus affect climate.	**Energy and Matter** • The total amount of energy and matter in closed systems is conserved.
Clarifications and Content Limits	**Clarification Statements** • Emphasis is on modeling biogeochemical cycles that include the cycling of carbon through the ocean, atmosphere, soil, and biosphere (including humans), providing the foundation for living organisms. **Content Limits** • <u>Students do not need to know:</u> How to calculate the residence time by dividing the reservoir size by the flow rate, either in or out; how to calculate the biomass in a given ecosystem.		
Science Vocabulary Students are Expected to Know	Concentration, rate of transfer/flow, pathway, hydrosphere, geosphere, biosphere, reservoir, sink, basin, pool, accumulate, biomass, equilibrium, chemosynthesis, byproduct, element, hydrocarbon, organic, inorganic, biotic, abiotic, diffusion, decompose, decay, microbe, fungi, bacteria, sediments, sequestered		
Science Vocabulary Students are Not Expected to Know	assimilation, residence time, facies, orogenic, strata, outgassing, LeChatelier's Principle		
Phenomena			
Context/ Phenomena	Some example phenomena for HS-ESS2-6: • Data indicates that higher levels of atmospheric carbon dioxide increase both carbon's input and release from the soil. • Even though trees take up carbon dioxide from the atmosphere, scientists find little carbon accumulation in the soil of a North Carolina forest. • Human activity releases more than 30 billion tons of carbon dioxide into the atmosphere per year. However, scientists estimate that Earth's soil releases roughly nine times more carbon dioxide into the atmosphere than all human activities combined.		
This Performance Expectation and associated Evidence Statements support the following Task Demands.			
Task Demands			
1. Select or identify from a collection of potential model components, mathematical variables, and/or mathematical operators, including distractors, the components, variables, and/or operators needed to mathematically and/or quantitatively model the phenomenon. Components and mathematical variables might include/represent organisms, spheres, molecules and/or elements, chemical, physical, and/or biological			

	processes, and reservoirs. Operators might include symbols for addition, subtraction, multiplication, division, etc.
2.	Assemble or complete, from a collection of potential model components, mathematical variables, and/or mathematical operators, an illustration or flow chart that is capable of mathematically and/or quantitatively representing how matter and energy are continuously transferred within and between organisms and their physical environment. This does not include labeling an existing diagram.
3.	Describe, select, or identify the mathematical and/or quantitative relationships among components of a model and/or mathematical variables that describe how matter and energy are continuously transferred within and between organisms and their physical environment.
4.	Manipulate the components of a mathematical and/or quantitative model to demonstrate the changes, properties, processes, and/or events that act to result in the phenomenon.
5.	Make predictions about the effects of changes in the rate at which materials or elements move from one reservoir or sphere to another. Predictions can be made by manipulating model components, mathematical variables, and/or mathematical formulas, completing illustrations, selecting from lists with distractors, or performing calculations given sufficient information to do so.
6.	Given mathematical and/or quantitative models or diagrams of how matter and energy are continuously transferred within and between organisms and their physical environment, identify the pathways of matter and/or energy transfer within an environment and how they change in each scenario OR identify the properties of the environment that cause changes in the transfer of matter and/or energy within that environment.
7.	Identify missing components, mathematical variables, mathematical and/or quantitative relationships, or other limitations of the mathematic and/or quantitative model.

<table>
<tr><td>Performance Expectation</td><td colspan="3">HS-ESS2-7
Construct an argument based on evidence about the simultaneous coevolution of Earth’s systems and life on Earth.</td></tr>
<tr><td>Dimensions</td><td>Engaging in Argument from Evidence
• Construct an oral and written argument or counter-arguments based on data and evidence.</td><td>ESS2.D: Weather and Climate
• Gradual atmospheric changes were due to plants and other organisms that captured carbon dioxide and released oxygen.

ESS2.E: Biogeology
• The many dynamic and delicate feedbacks between the biosphere and other Earth systems cause a continual co-evolution of Earth’s surface and the life that exists on it.</td><td>Stability and Change
• Much of science deals with constructing explanations of how things change and how they remain stable.</td></tr>
<tr><td>Clarifications and Content Limits</td><td colspan="3">Clarification Statements
• Emphasis is on the dynamic causes, effects, and feedbacks between the biosphere and the Earth’s systems, whereby geoscience factors control the evolution of life, which in turn continuously alters Earth’s surface.
• Examples include how photosynthetic life altered the atmosphere through the production of oxygen, which in turn increased weathering rates and allowed for the evolution of animal life; how microbial life on land increased the formation of soil, which in turn allowed for the evolution of land plants; and how the evolution of corals created reefs that altered patterns of erosion and deposition along coastlines and provided habitats for the evolution of new life forms.

Content Limits
• Assessment does not include a comprehensive understanding of the mechanisms of how the biosphere interacts with all of Earth’s other systems.</td></tr>
<tr><td>Science Vocabulary Students are Expected to Know</td><td colspan="3">Plate tectonics, rock formation, geologic evidence, ocean basin, radioactive, rock strata, time scale, continental boundary, ocean trench, sedimentation, continental shelf, crustal deformation, crustal plate movement, fracture zone, convection, atmospheric composition, groundwater, igneous rock, metamorphic rock, sedimentary rock, water cycle, landslide, deposition, greenhouse gas, mass wasting, molten rock, surface runoff</td></tr>
<tr><td>Science Vocabulary Students are Not Expected to Know</td><td colspan="3">Ecosystem services, Anthropocene, eutrophication, ecohydrology, geomorphology, heterogeneity</td></tr>
<tr><td colspan="4" align="center">Phenomena</td></tr>
<tr><td>Context/ Phenomena</td><td colspan="3">Some example phenomena for HS-ESS2-7:
• Eospermatopteris fossils (first trees) begin to appear in rocks dated 390 million years. Fossils of Tiktaalik (four legged fish), one of the earliest land animals, are found in the rock layers above Esopermatopteris.
• The appearance of cyanobacteria is recorded in fossils that formed roughly 3.5 billion years ago. Superior Type banded iron formed roughly 1.8 to 2.7 billion years ago. It is characterized by alternating red and gray layers of iron rich minerals and silica rich minerals.
• The Rhynie Chert beds in Aberdeenshire Scotland contain detailed fossils of early plants. Bryophyte fossils from about 500 million years ago, show small simple structured plants.</td></tr>
</table>

	Cooksonia pertoni fossils from about 430 million years ago show plants that were larger, spore bearing, and contained tissues that move water through the plant (vascular). • In 2016 two-thirds of the Northern portion of the Great Barrier Reef experienced severe bleaching. The Great Barrier Reef prior to this event, was made up of corals with a variety of bright colors that attracted a variety of marine life. In 2016, the coral turned completely white and few fish inhabit the area where bleaching has occurred.

This Performance Expectation and associated Evidence Statements support the following Task Demands.
Task Demands
1. Articulate, describe, illustrate, or select the relationships, interactions, and/or processes to be explained. This may entail sorting relevant from irrelevant information or features.
2. Express or complete a causal chain explaining how Earth's systems coevolved simultaneously with life on Earth. This may include indicating directions of causality in an incomplete model such as a flow chart or diagram, or completing cause and effect chains.
3. Identify and/or describe additional relevant evidence not provided that would support or clarify the explanation of the simultaneous coevolution of Earth's systems and life on Earth. This may entail sorting relevant from irrelevant information or features.
4. Construct or identify from a collection, including distractors, an explanation based on evidence that explains how Earth's systems coevolved simultaneously with life on Earth.*
5. Describe, identify, and/or select information and/or evidence needed to support an explanation. This may entail sorting relevant from irrelevant information or features.
6. Identify patterns or evidence in the data that support conclusions about the relationship between the evolution of life on Earth and Earth's systems.

*denotes those task demands which are deemed appropriate for use in stand-alone item development

<table>
<tr><td>Performance Expectation</td><td colspan="3">HS-ESS3-1
Construct an explanation based on evidence for how the availability of natural resources, occurrence of natural hazards, and changes in climate have influenced human activity.</td></tr>
<tr><td>Dimensions</td><td>Constructing Explanations and Designing Solutions
• Construct an explanation based on valid and reliable evidence obtained from a variety of sources (including students’ own investigations, models, theories, simulations, peer review) and the assumption that theories and laws that describe the natural world operate today as they did in the past and will continue to do so in the future.</td><td>ESS3.A: Natural Resources
• Resource availability has guided the development of human society.

ESS3.B: Natural Hazards
• Natural hazards and other geologic events have shaped the course of human history; [they] have significantly altered the sizes of human populations and have driven human migrations.</td><td>Cause and Effect
• Empirical evidence is required to differentiate between cause and correlation and make claims about specific causes and effects.</td></tr>
<tr><td>Clarifications and Content Limits</td><td colspan="3">Clarification Statements
• Examples of key natural resources include access to fresh water (such as rivers, lakes, and groundwater), regions of fertile soils such as river deltas, and high concentrations of minerals and fossil fuels.
• Examples of natural hazards can be from interior processes (such as volcanic eruptions and earthquakes), surface processes (such as tsunamis, mass wasting and soil erosion), and severe weather (such as hurricanes, floods, and droughts).
• Examples of the results of changes in climate that can affect populations or drive mass migrations include changes to sea level, regional patterns of temperature and precipitation, and the types of crops and livestock that can be raised.

Content Limits
• <u>Students do not need to know:</u> distribution of specific resources</td></tr>
<tr><td>Science Vocabulary Students Are Expected to Know</td><td colspan="3">Renewable, non-renewable, mitigation, economic cost.</td></tr>
<tr><td>Science Vocabulary Students Are Not Expected to Know</td><td colspan="3">Biome</td></tr>
<tr><td colspan="4">Phenomena</td></tr>
<tr><td>Context/ Phenomena</td><td colspan="3">Some example phenomena for HS-ESS3-1:
• In 2001, 85% of Australians lived within 50 km of the ocean.
• There are large solar power plants in the southern California desert. California solar power had a capacity of 18,296 MW in 2016. In the same year, New York State had a capacity of 927 MW.
• As many as 1.5 million inhabitants of Dhaka, Bangladesh, have moved there from villages near the Bay of Bengal.
• After the eruption of Mt. Vesuvius in 79 AD, the city of Pompeii was completely buried in volcanic ash. The city was never reoccupied and was lost for more than 1,500 years.</td></tr>
</table>

This Performance Expectation and associated Evidence Statements support the following Task Demands.	
Task Demands	
7. Articulate, describe, illustrate, or select the relationships, interactions, and/or processes to be explained. This may entail sorting relevant from irrelevant information or features.	
8. Express or complete a causal chain explaining how resource availability/natural hazards/climate change drive changes in human society/population/migration. This may include indicating directions of causality in an incomplete model such as a flow chart or diagram, or completing cause and effect chains.*	
9. Identify evidence supporting the inference of causation that is expressed in a causal chain.	
10. Use an explanation to predict the change in human /activity given a change in resource availability/natural hazards/climate.	
11. Describe, identify, and/or select information and/or evidence needed to support an explanation.	
12. Construct an explanation based on evidence that explains that the availability of natural resources/occurrence of natural hazards/changes in climate have influenced human activity.*	

*denotes those task demands which are deemed appropriate for use in stand-alone item development.

Performance Expectation	**HS-ESS3-2** Evaluate competing design solutions for developing, managing, and utilizing energy and mineral resources based on cost-benefit ratios.		
Dimensions	**Engaging in Argument from Evidence** • Evaluate competing design solutions to a real-world problem based on scientific ideas and principles, empirical evidence, and logical arguments regarding relevant factors (e.g., economic, societal, environmental, ethical, considerations).	**ESS3.A: Natural Resources** • All forms of energy production and other resource extraction have associated economic, social, environmental, and geopolitical costs and risks as well as benefits. New technologies and social regulations can change the balance of these factors. **ETS1.B: Developing Possible Solutions** • When evaluating solutions, it is important to take into account a range of constraints, including cost, safety, reliability, and aesthetics, and to consider social, cultural, and environmental impacts *(secondary)*	
Clarifications and Content Limits	**Clarification Statements**: • Emphasis is on the conservation, recycling, and reuse of resources (such as minerals and metals) where possible, and on minimizing impacts where it is not. • Examples include developing best practices for agricultural soil use, mining (for coal, tar sands, and oil shales), and pumping (for petroleum and natural gas). Science knowledge indicates what can happen in natural systems—not what should happen.		
Science Vocabulary Students are Expected to Know	Renewable, non-renewable, mitigation, economic cost, irreversible, reversible, exponential, logarithmic, basin, sustainability, recycle, reuse, species, societal, wetland, groundwater, metal, consumption, per-capita, stabilize, fossil fuel, mining, conservation, extract, agriculture, timber, fertile land, solar radiation, biotic, abiotic, depletion, extinction, economics, manufacturing, technology,		
Science Vocabulary Students are Not Expected to Know	Trigonometric, derivative, feedback, regulation, dynamic, aquifer, hydrothermal, geopolitical, oil shale, tar sand, urban planning, waste management, fragmentation		
Phenomena			
Context/ Phenomena	Some example phenomena for HS-ESS3-2: • There is a tower in the middle of North Dakota with flames shooting out the top of it. • In Pennsylvania, a match is struck next to a running water faucet and a large flame appears. • On the Yangtze River in China, blades of an underwater turbine turn and generate electricity. • In the desert of Oman, a farmer uses seawater to irrigate crops.		
Task Demands			
1. Articulate, describe, illustrate, or select the relationships, interactions, and/or processes to be explained. This may entail organizing, interpreting and analyzing data, making calculations, and sorting relevant from irrelevant information or features.			
2. Identify evidence that supports and/or does not support the success of competing design solutions for developing, managing, and utilizing energy and mineral resources based on cost-benefit ratios, societal needs for that resource, and associated environmental risks and benefits.			

3. Describe, select, or identify components of competing design solutions for developing, managing, and utilizing energy and mineral resources based on cost-benefit ratios supported by given evidence.
4. Evaluate the strengths of competing design solutions for developing, managing, and utilizing energy and mineral resources based on cost-benefit ratios, societal needs for that resources, and associated environmental risks and benefits.
5. Use an explanation of the design solutions for developing, managing, and utilizing energy and mineral resources to evaluate which design solution has the most preferred cost-benefit ratio.

*denotes those task demands which are deemed appropriate for use in stand-alone item development

<table>
<tr><td>Performance Expectation</td><td colspan="3">HS-ESS3-3
Create a computational simulation to illustrate the relationships among the management of natural resources, the sustainability of human populations, and biodiversity.</td></tr>
<tr><td>Dimensions</td><td>Using Mathematics and Computational Thinking
• Create a computational model or simulation of a phenomenon, designed device, process, or system.</td><td>ESS3.C: Human Impacts of Earth Systems
• The sustainability of human societies and the biodiversity that supports them require responsible management of natural resources.</td><td>Stability and Change
• Change and rates of change can be quantified and modeled over very short or very long periods of time. Some system changes are irreversible.</td></tr>
<tr><td>Clarifications and Content Limits</td><td colspan="3">Clarification Statements
• Examples of factors that affect the management of natural resources include the costs of resource extraction and waste management, per-capita consumption, and development of new technologies.
• Examples of factors that affect human sustainability include agricultural efficiency, levels of conservation, and urban planning.

Content Limits
• Assessment for computational simulations is limited to using provided multi-parameter programs or constructing simplified spreadsheet calculations.</td></tr>
<tr><td>Science Vocabulary Students are Expected to Know</td><td colspan="3">Biosphere, geosphere, hydrosphere, renewable, non-renewable, mitigation, economic cost, irreversible, reversible, exponential, logarithmic, basin, ecological, biome, recycle, reuse, mineral, societal, wetland, consumption, per-capita, mining, conservation, extract, agriculture, timber, fertile land, solar radiation, biotic, abiotic, depletion, extinction, manufacturing, technology</td></tr>
<tr><td>Science Vocabulary Students are Not Expected to Know</td><td colspan="3">Trigonometric, derivative, feedback, regulation, dynamic, aquifer, hydrothermal, geopolitical, oil shale, tar sand, urban planning, waste management, fragmentation</td></tr>
<tr><td colspan="4">Phenomena</td></tr>
<tr><td>Context/ Phenomena</td><td colspan="3">Some example phenomena for HS-ESS3-3:
• The number of birds and other wildlife in an area decreases by 30% after a shopping mall is built in northern California.
• Two 1,330 square-foot homes are side by side in northern California. One has six solar panels on the roof, and the other does not. During one month in June, the one with solar panels produces less carbon dioxide than the other house by174 kilograms.
• Beetles are present throughout a forest. Chemicals are sprayed at intervals needed to control the beetles on one acre. Fifty years later, this acre is the only part of the forest that has oak trees.
• Three species of fish, the Colorado squawfish, the roundtail chub, and the bonytail chub became extinct in the years immediately following construction of the Glen Canyon Dam in Colorado.</td></tr>
<tr><td colspan="4">This Performance Expectation and associated Evidence Statements support the following Task Demands.</td></tr>
<tr><td colspan="4">Task Demands</td></tr>
</table>

1. Use data to calculate or estimate the effect of an action or solution on natural resources, the sustainability of human populations, and/or biodiversity.
2. Illustrate, graph, or identify features or data that can be used to determine the effects of an action or solution on natural resources, the sustainability of human populations, and/or biodiversity.
3. Estimate or infer the effects of an action or solution that affects natural resources, the sustainability of human populations, and/or biodiversity.
4. Compile the data needed for an inference about the impacts of an action or solution on natural resources, the sustainability of human populations, and/or biodiversity. This can include sorting out the relevant data from the given information (or choosing relevant inputs for a simulation).
5. Using given information, select or identify the criteria against which the solution should be judged.
6. Using a simulator, test a proposed action or solution and evaluate the outcomes; may include proposing modifications to the action or solution.*****
7. Evaluate and/or critique models, simulations, or predictions in terms of identifiable limitations and whether or not they yield realistic results.

*In order to satisfy this PE, the student must use a simulator. Therefore, this task demand must always be used.

<table>
<tr><td>Performance Expectation</td><td colspan="3">HS-ESS3-4
Evaluate or refine a technological solution that reduces impacts of human activities on natural systems.</td></tr>
<tr><td>Dimensions</td><td>Constructing Explanations and Designing Solutions
• Design or refine a solution to a complex real-world problem, based on scientific knowledge, student-generated sources of evidence, prioritized criteria, and tradeoff considerations.</td><td>ESS3.C: Human Impacts on Earth Systems
• Scientists and engineers can make major contributions by developing technologies that produce less pollution and waste and that preclude ecosystem degradation.

ETS1.B Developing Possible Solutions
• When evaluating solutions, it is important to take into account a range of constraints, including cost, safety, reliability, and aesthetics, and to consider social, cultural, and environmental impacts (secondary)</td><td>Stability and Change
• Feedback (negative or positive) can stabilize or destabilize a system.</td></tr>
<tr><td>Clarifications and Content Limits</td><td colspan="3">Clarification Statements
• Examples of data on the impacts of human activities could include the quantities and types of pollutants released, changes to biomass and species diversity, or areal changes in land surface use (such as for urban development, agriculture and livestock, or surface mining).
• Examples for limiting future impacts could range from local efforts (such as reducing, reusing, and recycling resources) to large-scale geoengineering design solutions (such as altering global temperatures by making large changes to the atmosphere or ocean).</td></tr>
<tr><td>Science Vocabulary Students are Expected to Know</td><td colspan="3">Renewable, non-renewable, mitigation, economic cost, irreversible, reversible, exponential, logarithmic, basin, recycle, reuse, societal, wetland, metal, consumption, per-capita, biodiversity, stabilize, mining, conservation, extract, agriculture, timber, fertile land, solar radiation, biotic, abiotic, depletion, extinction, economics, manufacturing, technology</td></tr>
<tr><td>Science Vocabulary Students are Not Expected to Know</td><td colspan="3">Trigonometric, derivative, feedback, regulation, dynamic, aquifer, hydrothermal, geopolitical, oil shale, tar sand, urban planning, waste management, fragmentation</td></tr>
<tr><td colspan="4" align="center">Phenomena</td></tr>
<tr><td>Context/ Phenomena</td><td colspan="3">Some example phenomena for HS-ESS3-4:
• Recycling and composting almost 87 million tons of municipal solid waste saved more than 1.1 quadrillion Btu of energy; roughly equivalent to the same amount of energy consumed by 10 million U.S. households in a year.
• Mixed Paper recycling saves the equivalent of 165 gallons of gasoline.</td></tr>
<tr><td colspan="4" align="center">This Performance Expectation and associated Evidence Statements support the following Task Demands.</td></tr>
<tr><td colspan="4" align="center">Task Demands</td></tr>
<tr><td colspan="4">1. Articulate, describe, illustrate, or select the relationships, interactions, and/or processes to be explained. This may entail organizing, interpreting and analyzing data, making calculations, and sorting relevant from irrelevant information or features.</td></tr>
<tr><td colspan="4">2. Identify evidence that supports and/or does not support the success of the technological solution that reduced impacts of human activities on natural systems.</td></tr>
</table>

3. Describe, select, or identify components of the impacts of human activities on natural systems supported by given evidence.
4. Use an explanation of the impacts of human activities on natural systems to explain the technological solution.
5. Identify or select the information needed to support an explanation of the impacts of human activities on natural systems.
6. Using given information about the effects of human activities on natural systems, select or identify criteria against which the solution should be judged.
7. Using given information about the effects of human activities on natural systems, select or identify constraints that the solution must meet.
8. Evaluate the criteria and constraints, along with trade-offs, for a proposed or given solution to resolve or improve the impact of human activities on natural systems.
9. Using given data, propose a potential solution to resolve or improve the impact of human activities on natural systems.
10. Using a simulator, test a proposed solution to resolve or improve the impact of human activities on natural systems, biodiversity and evaluate the outcomes.
11. Evaluate and/or revise a solution to resolve or improve the impact of human activities on natural systems, and evaluate the outcomes

<table>
<tr><td>Performance Expectation</td><td colspan="3">HS-ESS-3-5
Analyze geoscience data and the results from global climate models to make an evidence-based forecast of the current rate of global or regional climate change and associated future impacts on Earth’s systems.</td></tr>
<tr><td>Dimensions</td><td>Analyzing and Interpreting Data
• Analyze data using computational models in order to make valid and reliable scientific claims.</td><td>ESS3.D: Global Climate Change
• Though the magnitudes of human impacts are greater than they have ever been, so too are human abilities to model, predict, and manage current and future impacts.</td><td>Stability and Change
• Change and rates of change can be quantified and modeled over very short or very long periods of time. Some system changes are irreversible.</td></tr>
<tr><td>Clarifications and Content Limits</td><td colspan="3">Clarification Statements
• Examples of evidence, for both data and climate model outputs, are for climate changes (such as precipitation and temperature) and their associated impacts (such as sea level, glacial ice volumes, or atmosphere and ocean composition).

Content Limits
• Assessment is limited to one example of a climate change and its associated impacts.</td></tr>
<tr><td>Science Vocabulary Students are Expected to Know</td><td colspan="3">Orientation, probabilistic, redistribute, volcanic ash, concentration, electromagnetic radiation, radiation, sea level, geochemical reaction, geoscience, greenhouse gas, atmospheric change, biosphere, global temperature, ice core, methane, glacier</td></tr>
<tr><td>Science Vocabulary Students are Not Expected to Know</td><td colspan="3">Anthropogenic, absorption spectrum, determinant, NOX, Carbon Footprint,</td></tr>
<tr><td colspan="4">Phenomena</td></tr>
<tr><td>Context/ Phenomena</td><td colspan="3">Some example phenomena for HS-ESS3-5:
• The model predictions for the Great Lakes region of the United States consist of increased precipitation of 5-30% during the spring and decreased precipitation of 5-10% in the summer.
• Concentrations of CO_2 under the higher emissions scenario for 2100 could reach as high as 850 parts per million (ppm).
• Global warming of 2°C is predicted by the year 2050
• The model mean global temperature change for a high emissions scenario is 4-6°</td></tr>
<tr><td colspan="4">This Performance Expectation and associated Evidence Statements support the following Task Demands.</td></tr>
<tr><td colspan="4">Task Demands</td></tr>
<tr><td colspan="4">1. Organize and/or arrange (e.g., using illustrations and/or labels), or summarize data to highlight trends, patterns, or correlations in global or regional climate models and their associated future impacts on Earth’s systems.</td></tr>
<tr><td colspan="4">2. Generate/construct graphs, tables, or assemblages of illustrations and/or labels of data that document patterns, trends, or correlations in global or regional climate models to forecast regional climate change and the associated future impacts on Earth’s systems. This may include sorting out distractors.</td></tr>
<tr><td colspan="4">3. Use relationships identified in the data to forecast the current rate of global or regional climate change and how it will affect Earth’s systems.</td></tr>
</table>

4. Identify patterns or evidence in the data that supports inferences about how the changing of global or regional climates will affect Earth's systems in the long term.

This page was intentionally left blank.

Performance Expectation	**HS-ESS-3-6** Use a computational representation to illustrate the relationships among Earth systems and how those relationships are being modified due to human activity.
Dimensions	**Using Mathematics and Computational Thinking** • Use a computational representation of phenomena or design solutions to describe and/or support claims and/or explanations. **ESS2.D: Weather and Climate** • Current models predict that, although future regional climate changes will be complex and varied, average global temperatures will continue to rise. The outcomes predicted by global climate models strongly depend on the amounts of human-generated greenhouse gases added to the atmosphere each year and by the ways in which these gases are absorbed by the ocean and biosphere. *(secondary)* **ESS3.D: Global Climate Change** • Through computer simulations and other studies, important discoveries are still being made about how the ocean, the atmosphere, and the biosphere interact and are modified in response to human activities. **Systems and System Models** • When investigating or describing a system, the boundaries and initial conditions of the system need to be defined and their inputs and outputs analyzed and described using models.
Clarifications and Content Limits	**Clarification Statements** • Examples of Earth systems to be considered are the hydrosphere, atmosphere, cryosphere, geosphere, and/or biosphere. • An example of the far-reaching impacts from a human activity is how an increase in atmospheric carbon dioxide results in an increase in photosynthetic biomass on land and an increase in ocean acidification, with resulting impacts on sea organism health and marine populations. **Content Limits** • Assessment does not include running computational representations but is limited to using the published results of scientific computational models.
Science Vocabulary Students are Expected to Know	Orientation, probabilistic, redistribute, volcanic ash, concentration, electromagnetic radiation, radiation, sea level, geochemical reaction, geoscience, greenhouse gas, atmospheric change, biosphere, global temperature, ice core, methane, glacier
Science Vocabulary Students are Not Expected to Know	Anthropogenic, absorption spectrum, determinant, NOX, Carbon Footprint,
Phenomena	
Context/ Phenomena	Some example phenomena for HS-ESS3-6: • Beetles are present throughout a forest. Chemicals are sprayed at intervals needed to control the beetles on one acre. Fifty years later, this acre is the only part of the forest that has oak trees. • In July 2016, the size of the hypoxic area due to algae blooms in the Chesapeake Bay in late June was the second smallest since 1985.
This Performance Expectation and associated Evidence Statements support the following Task Demands.	

Task Demands
1. Use data to calculate or estimate the effect of human activity on Earth systems.
2. Illustrate, graph, or identify features or data that can be used to determine the relationships among Earth systems and how human activity is affecting those relationships.
3. Estimate or infer the effects of human activity on Earth systems.
4. Compile the data needed for an inference about the impacts of human activity on Earth systems. This can include sorting out the relevant data from the given information (or choosing relevant inputs for a simulation).
5. Using a simulator, test a prediction and evaluate the outcomes. This may include proposing modifications to the action to mitigate or the solution to the effect(s) of human activity on Earth systems.
6. Evaluate and/or critique models, simulations, or predictions in terms of identifiable limitations and whether or not they yield realistic results.

Appendix 2-C

Style Guide for Science Items

Cambium Assessment
(formerly AIR Assessment)

Three-Dimensional Science Assessment Style Guide

Modified from the Smarter Balanced Style Guide

3-11-2022

Note: The presentation of the sample items and selections in this document approximates but does not exactly reflect the appearance of the test content that students will view on the computer screen. The final presentation of content will depend on the user interface (UI) of the online delivery system.

Table 1. Abbreviations used in Style Guide

Style Guide Abbreviations	
Abbreviation	Spelled-Out Term
CBT	computer-based testing
CMOS	*Chicago Manual of Style*
CMYK	cyan-magenta-yellow-black (a four-color model used in printing)
CR item	constructed-response item
Dpi	dots per inch
JPEG	Joint Photographic Experts Group (a format for compressing images used for print or screen
PBT	print-based testing
PNG	portable network graphics (a format using lossless compression for images used for screen presentation)
RGB	red-green-blue (a three-color model used in screen presentation)
SR item	selected-response item
TEI	technology-enhanced item
TIFF	tagged image file format (a format for compressing images used for print presentation)
UI	user interface

Table of Contents

Part I: Global Style Conventions

A. Computer-Based Testing

Computer-based testing (CBT) differs from traditional print-based testing (PBT). Traditional style conventions applied to printed test forms must be modified for computer-based test forms because not all print-based style conventions are appropriate for display on a computer screen.

For example, the Verdana font has been chosen because it was specifically designed to be used in place of Times New Roman and other serif fonts that often appear in printed test forms. Verdana characters are slightly larger than characters in other fonts, and the ample space between the characters makes them easy to distinguish at low screen resolutions.

Layout Considerations

The presentation of content in computer-based test forms depends on the user interface (UI) in the online delivery system. The guidelines in this section should be applied to the extent possible once the online delivery system is identified.

Content panes

Students should have the option of viewing content in one pane that is the full size of the computer screen or in sub-panes that divide the screen horizontally or vertically. For example, students should be able to view a selection on the screen by itself or on the screen with an item.

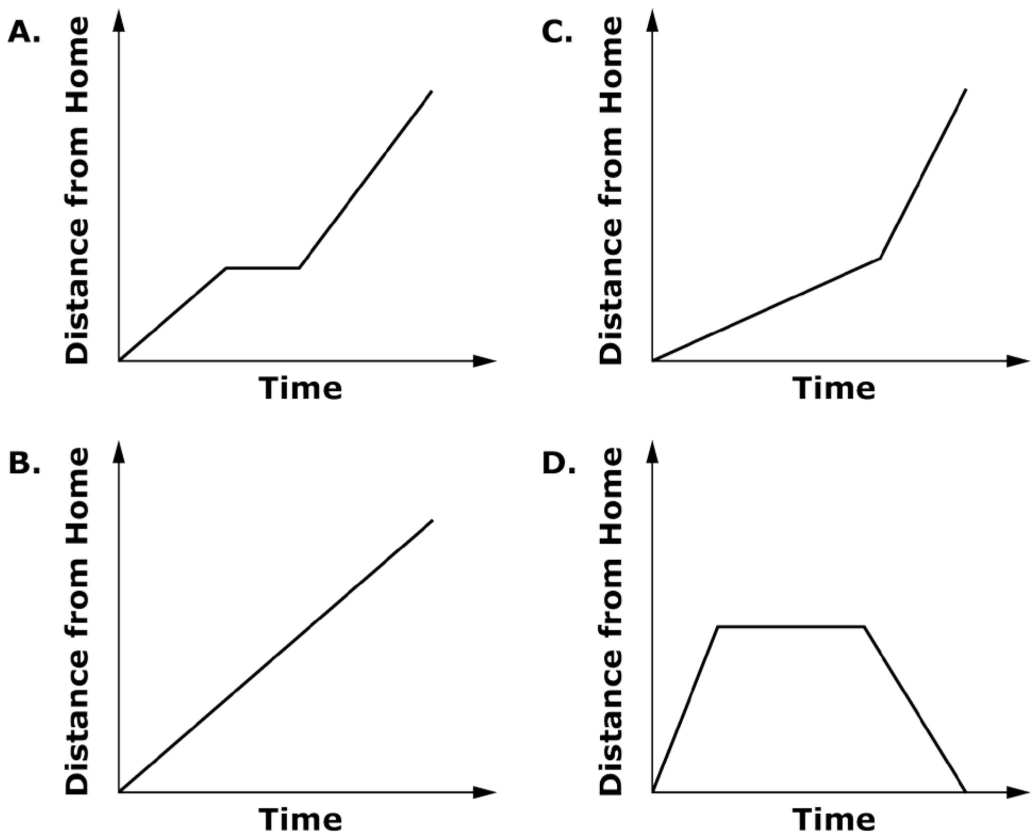
The sizes of content panes depend on the amount of space allotted in the UI for viewing content. For example, content panes will be larger in a UI in which 80% of the viewing area is used for content and 20% is reserved for headers, navigation tools, and other non-content elements than in a UI in which 70% of the viewing area is used for content and 30% is reserved for non-content elements.

Scrolling

Students should not have to use horizontal scrolling to view test content in its entirety. The following guidelines are designed to minimize the need for scrolling:

- Lay out content across the computer screen rather than in columns.
- Display each item on the screen by itself.
- Allow students to view a selection on the screen by itself or on the screen with an item. This eliminates the electronic equivalent of turning a page to flip between the selection and the items.
- Similar options should be provided for viewing a cluster. Students should be able to view the graphic on the screen by itself or on the screen with an item in the cluster.
- Whenever possible, arrange graphic options in a two-over-two box (stacked) layout.

Which graph shows that Noel walked home from school at a constant rate?



Line Breaks

The locations of line breaks depend on the operating system on the computer as well as the size of the monitor being used to view test content. Hard line breaks should not be inserted unless absolutely necessary because lines of text do not necessarily break at the same locations from computer to computer.

B. General Font and Alignment Specifications

Table 2 shows general font and alignment specifications for different text elements. See "Selected-Response Items" in this section for information about option alignment.

Note: All text should be displayed on a white background.

Table 2. General font and alignment specifications for text elements in test forms

Content	Font	Alignment
Items	14 pt. Verdana	Left aligned
Part headings in items	14 pt. Verdana Bold	Left aligned
Boxed text	14 pt. Verdana	• Box: left aligned • Text: longest line centered in box; other lines left aligned on longest line.
Emphasis terms	14 pt. Verdana Bold	n/a

Note: At this time, Verdana is specified as the primary font for test content. However, another font may be chosen upon further analysis of the effects that fonts have on readability and students' ability to retain information.

C. Stimulus Specifications

The stimulus of a cluster describes a scientific phenomenon and provides the student with the necessary context around that phenomenon to answer the associated items. A stimulus may include a title, but a title is not necessary.

Phenomenon
The phenomenon is a scientific observation or an engineering problem around which the cluster is based. The phenomenon should be stated as the first sentence(s) of the stimulus, followed by two hard returns.

Example phenomenon:

> Fog appears and disappears over the course of a morning in the Willamette Valley in Oregon.

Task Statement
The task statement concludes the stimulus by telling the student exactly what he or she is expected to do in the cluster. The task statement is preceded with "Your Task" in bold, followed by two hard returns.

Example task statement:

> **Your Task**
>
> In the questions that follow, you will develop an explanation for the appearance and disappearance of fog.

Note: A standalone item with a stimulus does not have a task statement ("Your Task").

Stimulus Layout
If the stimulus contains information that students will consistently reference when working through the items in the cluster, the writer should select a split-screen layout that allows the stimulus to appear on the left and the items on the right. If the stimulus only serves as an introduction to the phenomenon, the writer should select a single screen layout. Animations, graphics, and other stimulus material should be sized to avoid the need for horizontal scrolling. If it becomes difficult to limit the size of an animation or another element of the stimulus, a single screen layout may be selected to avoid horizontal scrolling.

Wording should be consistent throughout the stimulus and cluster. For example, do not use "steam" and "vapor" interchangeably.

D. Interaction Types

Science assessments consist of various types of interactions, including selected-response (SR) items, constructed-response (CR) items, and technology-enhanced items (TEIs). In test forms, items are numbered sequentially, beginning with number 1.

This section provides general global style conventions and specifications for items. See Part II for content specific conventions and for information about TEIs.

Selected Response Interactions
All SR interactions consist of a stem and options. The format of the stem and options varies among items based on content. (See "Options" in this section for formats of options.)

Stems:

- Stem: the part of the SR item that precedes the options
- Closed stem: a stem that is a complete sentence and ends with a period or question mark
- Open stem: a stem that consists of a sentence fragment and becomes a complete sentence when combined with each option.

Closed stems
When the interaction stem is closed, the options are either complete sentences or sentence fragments. Options that are complete sentences begin with a capital letter and end with a punctuation mark.

How would a fish population affect the stream ecosystem?
A. Fish would lower the water temperature.
B. Fish would produce oxygen from the water.
C. Fish would block sunlight, increasing plant growth.
D. Fish would produce waste, providing nutrients to plants.

Options that are fragments begin with a lowercase letter (unless the first word is a proper noun or adjective) and do not end with a punctuation mark.

The traits of populations in the forest ecosystem have changed over time. What caused the traits to change?
A. natural selection
B. lack of mutations
C. unlimited resources
D. asexual reproduction

The treatment of options that are imperative sentences depends on whether the implied subject of the sentence in each option is "you." If it is, the options are treated as complete sentences. If it is not, the options are treated as fragments.

Subject is implied "you":
What should you do next in the experiment?

A. Water the plants.
B. Label the volumes.
C. Cut the plant stems.
D. Record plant heights.

Subject is not implied "you":
What should Fiona do next in the experiment?

A. water the plants
B. label the volumes
C. cut the plant stems
D. record plant heights

Open Stems
When the item stem is open, both the stem and the options are fragments that, when combined, form complete sentences. The fragment in the stem does not end with a punctuation mark. Regardless of whether the options are complete sentences or fragments, they begin with a lowercase letter (unless the first word is a proper noun or adjective) and end with a punctuation mark.

The independent variable of the investigation is
A. volume.
B. height.
C. mass.
D. time.

In open-stem items, the options should not repeat large quantities of text. The stem must be long enough to provide context for the options. There is **no** punctuation at the end of an open stem (i.e., dash or colon).

The tussock moths obtain energy in cellular respiration by
A. taking in water.
B. releasing oxygen.
C. breaking down glucose.
D. inhaling carbon dioxide.

Options
Although all SR items have options, the number and format of these options vary from item to item based on content. The number of correct answers among the options also varies.

- In the case of a multiple choice interaction, options are identified with consecutive uppercase letters.

What is one purpose of ATP molecules in plant and animal cells?
A. to increase the rate of diffusion across cell membranes
B. to decrease the rate of chemical reactions
C. to store energy used for cell processes
D. to pass genetic traits to offspring

- In the case of a multiple select interaction, options are preceded by open boxes, which students can click on to select the correct answers.

Select **all** the cell structures found in animal cells.

- ◻ cell membrane
- ◻ mitochondria
- ◻ chloroplast
- ◻ cell wall
- ◻ nucleus

Option alignment and order

Table 3 provides general guidelines for the alignment and order of options in SR items.

Note: Options derived from a stimulus, such as a selection or graphic, are ALWAYS arranged in the same order in which they appear in the stimulus. This guideline supersedes all other guidelines listed in Table 3.

Table 3. Guidelines for the alignment and order of options in SR items

Option Alignment and Order			
Option Type	**Alignment**	**Order**	**Example**
Graphic options	• Graphic: left aligned • Option letter: top aligned or vertically centered on graphic (use best judgment)	Arranged for best visual presentation (use best judgment)	n/a
Numeric options	Decimal aligned: • stand-alone numbers • decimal values • numbers that precede or follow symbols: 40°, $20.00 • numbers that precede labels: 6 ties, 12 bananas • numbers that precede units of measure: 15 kilograms, 30 cm	Arranged in ascending or descending order	What percentage of students prefer strawberry yogurt to blueberry yogurt? A. 25% B. 50% How long, in centimeters, is each necklace? A. 9 cm B. 12 cm
	Fractions: • Fraction: left aligned • Option letter: vertically centered on fraction	Arranged in ascending or descending order	How many cups of sugar does the student need to make two cakes? A. 1/2 B. 1/4
	Times of day: • Decimal aligned on colon	Arranged in ascending or descending order	At what time does the student eat lunch? A. 11:30 a.m. B. 1:00 p.m.

Table 3. Guidelines for the alignment and order of options in SR items *(cont.)*

Option Type	Alignment	Order	Example
Text options	Left aligned	Words: arranged in ascending or descending order by word length	What large molecule is made of many small amino acid molecules? A. lipid B. protein C. carbohydrate
	Left aligned	Phrases and sentences: arranged by length, longest to shortest or vice versa; if more than one line of text is used as separate paragraphs (such as a title followed by a description), the length of the first line of text is considered (e.g., the title)	Which statement describes DNA? A. It is a macromolecule. B. It is found in the nucleus. C. It makes up chromosomes.

Note: Options are not arranged in the prescribed order when doing so cues the answer to the item.

Constructed-Response Interactions

CR items consist only of item stems. The stems are complete sentences written as imperative commands.

Explain how the amount of sunlight affects the growth of plants.
Use information from Investigation 1 to support your answer.

Describe how meteorologists predict the weather.
Use information from the Weather Patterns map in your description.

Bullets in Items

- Bullets are used for student action items—what a student must do to earn points, as listed in the rubric.
- Do not use a bullet if there is only one directive in an item.
- Always use bullets to denote multiple directives.

Parts in Items

Some items are divided into parts labeled with consecutive lettered headings that are followed by two hard returns. In headings, the word *part* and the part letter are capitalized and bold. In item text, the word *part* is lowercase and the part letter is capitalized. Parts may comprise a mix of interaction types or multiple interactions of the same interaction type.

Part A

Select a testable, scientific question that can be answered by performing an experiment with the setup shown in the Hanging Magnets Experiment picture.

Ⓐ How does the distance between the magnets affect the force?

Ⓑ How does the orientation of the magnets affect the force?

Ⓒ Will the force between the magnets always exist?

Part B

Use the table to select the properties you want to hold constant and the properties you want to change when you run your experiment to answer the question you chose in part A.

E. Graphics in Items

Introductory Statements

Descriptive terms
A graphic in an item should not be referred to as a *graphic*; instead, use a more descriptive term, such as *graph, table, or diagram*. The graph, table, or diagram should be labeled with a Figure or Table number. Any references in the stimulus or items should use this label rather than the title.

> Figure 1. Bird's Nest in Tree
> Use Figure 1 to describe …
>
> Table 1. Population of Zebra in the Serengeti
> Use Table 1 to predict …

Above and below
Do not use the terms *above* and *below* to refer to the location of a graphic; instead, use terms such as *this, that, an,* and *the*.

> An equation is shown.
>
> $$\square + 12 = 42$$
>
> Select the variables that make the equation true.

Introductory statements in the item stems
Graphics positioned within items are aligned to the **center** and introduced as part of the item stem using the Figure or Table number.

> Figure 1 shows a diagram of photosynthesis.
> Table 1 shows characteristics of two populations of coyote.

An introductory statement that is part of an item stem should be as descriptive as possible; however, a sentence such as "See Figure 3" is acceptable when a more descriptive introduction is not available or appropriate.

> *Acceptable:*
> See Figure 3.
>
> *Preferred:*
> Figure 3 shows food relationships among some organisms in the forest ecosystem.

Note: In SR items, stem text will be used before the graphic to introduce the item. The graphic will appear next. The graphic will be followed by any remaining stem text, including the question part of the stem. The options will appear last.

Referring to Text from Graphics in Items

Items sometimes refer to table column headings, labels, and other text in graphics. Use the guidelines that follow to determine the appropriate treatment of text from graphics that is referred to in items. Note that these guidelines apply only to the treatment of graphic text in items, not to the text in the graphic itself. (See Part III for information about the treatment of text in graphics.)

- Text from graphics is not enclosed in quotation marks.
- Titles and headings appear in regular type and follow the same capitalization used in the graphic.

  Which unit should be used for the measurements in the column titled Volume of Water?

- Labels and other text appear in regular type but do not necessarily follow the same capitalization used in the graphic. Use the guidelines that follow to determine whether to capitalize labels and other text from graphics.
- Capitalize labels that precede letters or numbers if the label is capitalized in the graphic.

  Data from the experiment are shown.

Trial	**Speed (m/s)**
1	10
2	15
3	25
4	20

  Why did Trial 1 have an outcome that was different from the outcome in Trial 3?

- Capitalize labels that are proper nouns.
- Lowercase labels that are common nouns.

  Jesse walks three dogs in his neighborhood. Chart 1 shows the number of times he walked each dog last week.

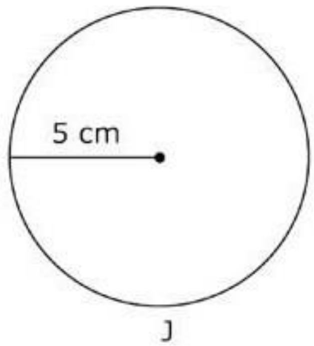
Chart 1

Dog	**Number of Times Walked**
Schnauzer	5
Beagle	4
Labrador retriever	10

  How many more times did Jesse walk the Labrador retriever than he walked the schnauzer?

- Elements are not capitalized when written out (i.e. hydrogen, oxygen)
- Lowercase labels that are lowercase or not shown in the graphic.

The radius of circle J is 5 centimeters.



The radius of circle K is twice as long as the radius of circle J. What is the **diameter**, in centimeters, of circle K?

F. Excerpts and Citations

Excerpts and citations for external sources should be formatted as shown below. Use quotation marks for titles of articles; italicize titles of longer works such as periodicals, journals, and books; and underline web addresses, but make sure there is no live link.

Source 1: Excerpt from "Light as a Broad-Spectrum Antimicrobial" by Peter Gwynne and Maurice Gallagher. Published in the academic journal *Frontiers in Microbiology*, 2018 (adapted from original).

> Research exists regarding the use of light as a sterilizing agent in food and water processing, much of it based around UV light. UV, however, penetrates poorly, limiting its application to surface decolonization. Although UV light is powerful, visible light wavelengths may be better in application.

Source 2: Excerpt from "Ultraviolet Disinfection of Drinking Water" by the Government of Western Australia Department of Heath. Published on the website https://ww2.health.wa.gov.au, 2016 (adapted from original).

> UV light will only travel in a straight line, so any obstruction will reduce its efficiency. Water that has not been filtered can contain iron, manganese, and other particles that can either absorb or scatter UV light, reducing the effectiveness of the system. Bacteria that are able to pass through—protected by shadows created by dirt, debris, or other particles—may be able to survive treatment.

G. Scoring Assertion Specifications

Item writers create scoring assertions to outline the criteria used to score the item. Scoring assertions should capture the features of the student response that receive credit *and* the inference that the test developer would like to make from that evidence. The example below shows several scoring assertions for a multi-interaction item. For each assertion, the test developer describes the features of the student response that receive credit ("The student [selected, identified, etc.] …), then links that feature to an inference about student understanding ("providing evidence of [the ability to, an understanding of, etc.] ….").

Score Rationale	
The student identified distance as affecting the force, providing evidence of the ability to form a conclusion based on experiment data.	✓
The student ran three trials keeping the weight of the box the same, providing evidence of an understanding of how to control variables in an investigation.	✓

Each cluster must have the **minimum** number of assertions when it goes into the Locked Operational Pool. To accomplish this, clusters should be created and revised to have at least two more assertions than the minimum to account for interactions and assertions being changed, collapsed, or rejected throughout the review process.

- **Minimum** number of assertions for final operational version of a cluster:
  - Elementary: 5
  - Middle: 7
  - High: 7

- **Recommended minimum** number of assertions for working drafts/versions of a cluster:
  - Elementary: 7
  - Middle: 9
  - High: 9

H. Exemplar Conventions

When to Create Exemplars

There are three main reasons to create exemplars.

1. There must be an exemplar created for each point value before an item can be web approved.
    a. Considering that clusters can have a larger number of possible point combinations, we propose that each exemplar created after the first exemplar adds the next consecutive assertion. See example below.
2. There should be an exemplar for each dependent scoring possibility, or one example in the case of infinite possibilities.
    a. The dependent assertion exemplar should highlight the specific dependency. We recommend that each exemplar have the minimum number of assertions marked as true (only respond to the parts relevant to the dependency), so that it is easier to highlight the specific assertions and interactions involved in the dependency. This will likely mean that most dependency exemplars will be 1 point.
3. Exemplars can be created to highlight responses that are not scoring properly.
    a. The reviewer should leave a comment in ITS that explains why the exemplar was created.

        "The dependent scoring in part B does not work as described in the assertion text. See CR1Review_PartB for the response that should receive credit but does not."

    b. The exemplar should be deleted when the issue is resolved.

Naming Conventions

Each exemplar should begin with a tag that indicates what type of exemplar it is, followed by a short description.

- Standard
    - This denotes "normal" scoring.
    - Follow with the point value:
        - Standard_2pt
        - Standard_FullCredit
- Dependency
    - This denotes that the exemplar relates to dependent scoring.
    - Follow with what part or parts are involved in the scoring, as well as a short description if there are several dependencies in those same parts.
        - Dependency_PartAB
        - Dependency_PartABC_greater
        - Dependency_PartA
- Review
    - This denotes a response that a reviewer had specific concerns about.
    - Begin with the review level in which the issue was found.
        - CR1Review_1pt
        - CR1Review_PartABC_greater

Example from Staging Site – IAT Sandbox 18402

Name	Points	Actions
Standard_FullCredit	3	View Delete
Standard_2pt	2	View Delete
Standard_1pt	1	View Delete
Dependency_PartAB	1	View Delete
Dependency_PartABC_greater	1	View Delete
Dependency_PartABC_equal	1	View Delete
CR1Review_1pt	1	View Delete
No response	0	View Delete

Exemplar: Standard_2pt

Saved By: Matthew_Davis
Date Saved: 4/22/2020 2:04:25 PM
Approved: No

Score response when exemplar was saved:
Your respnse earned **2** points of a possible **3**

Score Rational	
The student entered a value of 5.	✔
The student entered a value of 7 for y, or a value that corresponds to the value of x entered in part A.	✔
The student selected x < y or a selection corresponding to their values in parts A and B.	✖

I. Option Rationales (Formative, specific projects only)

IAT Options

- Multiple Choice: Select only one option.
- Multiple Select: Select two or more options.

Formatting Rationales

- Rationales for correct answers start with: "Key - …"
- All rationales provide a brief reasoning for why the response is a key or distractor relative to the prompt.

Multiple Choice Example

How does energy flow through the ecosystem?

A. The algae produce their own food.
Key - The algae are producers and, therefore, make their own food.

B. The salmon makes its own energy.
The salmon is a consumer and must get its energy from eating other organisms.

C. The sculpin relies on the salmon for energy.
The sculpin is eaten by the salmon.

D. The mollusk produces energy from the salmon.
The mollusk is not eaten by the salmon.

Multiple Select Example

Select the **three** fossils of organisms that lived more than 300 mya (million years ago).

□ Fossil 1
This fossil is found in the rock layer at the top of the column, which means the organism lived less than 30 mya.

□ Fossil 2
Key - This fossil is found in the second rock layer from the bottom of the column, which means that the organism lived 450–500 mya.

□ Fossil 3
Key - This fossil is found in the bottom rock layer of the column, which means the organism lived over 550 mya.

□ Fossil 4
This fossil is found in the second rock layer from the top of the column, which means the organism lived less than 100 mya.

□ Fossil 5
Key - This fossil is found in the bottom of the rock layer of the column, which means the organism lived over 550 mya.

J. Paper Style

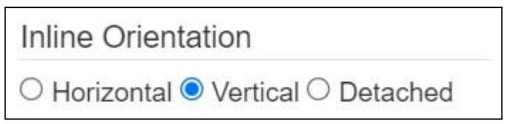
Compatible Item Types

ETC, MC, MS, Table Match, EQ

In most cases, language should be edited as little as possible to make the directions appropriate for paper use. However, the goal for the students is to make the testing experience as similar to the online experience as possible, so in some cases, the addition of directions is recommended in order to clarify instructions for interactions that are clearer online.

ETC Interactions

- Instead of "Click on the blank boxes to select …" stems should read "Select the word or phrase …" or "Select a word or phrase …."
  - If there are multiple ETCs in one interaction, use "Select a word or phrase in each box to …" or similar. It needs to be clear that only one selection in each grouping should be made. In causal chains especially (depending on the style), it needs to be clear that only one statement should be selected for each step. In these cases, "in each box" should be included in the stem for clarity.
  - In the ETC interaction, change the options to orient vertically. They should appear vertically in a box rather than next to each other horizontally.



- For ranking interactions, use "Fill in the bubbles to rank …."
- For interactions in which it is not obvious that students should make only one selection for each step—such as flowcharts or tables—add the statement "Make only one selection for each step." "Step" may be replaced with an appropriate term for the interaction.
  - This statement is added because online, once students make a selection, that is the only selection shown, so they cannot accidentally select another choice. On paper, students may accidentally fill in multiple bubbles in the same step. The addition of this wording reduces the likelihood that students would make this mistake, therefore making the experience closer to the online testing experience.

MC and MS Interactions

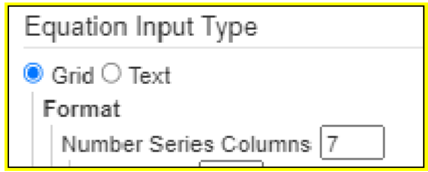
- Wording should match online wording.
- For choice interactions with graphics as options that are all vertical, consider changing the options to 2x2 (two rows with two options each).

Table Match Interactions

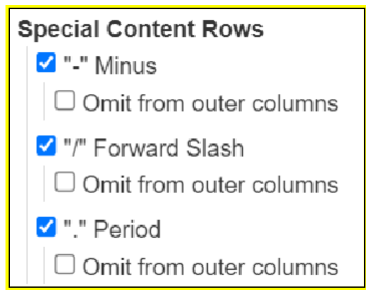
- Wording should be changed to "Fill in the bubbles to select/identify/etc. …."
- If the online version of a table match allows only one box to be checked per row or column and this would not be obvious to the student viewing the item on paper, add verbiage to make this clearer, e.g., "Fill in one bubble in each row."

Equation Interactions

- Check that the stem makes sense, given the appearance of the EQ on paper, and edit if needed. Include "Enter a number in each blank box." In the EQ Paper Renderer, always include 7 digits (columns), for **all** grade levels. This is consistent with the Math team's paper style.



- Make the bubbles match what is available in the online rendering as much as possible. That is, if subtraction, decimal, and/or fraction/division is allowed in the online version, include these for paper as well (note, however, that multiplication and addition are not available for paper). If they are not allowed for online, do not include them for paper. Machine scoring for paper works the same way as the scoring in the grub, so we can allow students to enter expressions on paper the same way we do for online.



Graphics

All graphics will be converted to grayscale. Ensure that graphics render appropriately. Graphics that require students to differentiate between colors should be avoided when possible. If the item refers to the colors in stems or options, but colors are not labeled in the graphics, labels should be added. Check any color graphics with the monochromacy view on the Coblis website (https://www.color-blindness.com/coblis-color-blindness-simulator/) to see whether edits are needed (e.g., adding patterns if colors appear the same in grayscale).

Example: Figure 1 shows parent flowers and offspring of various colors, and those colors are not labeled in the graphic. The question relating to Figure 1 is written as:

Describe the inheritance pattern of the flowers shown in Figure 1.
A. Yellow flowers inherit their color from their parents.
B. Orange flowers inherit their color from their parents.
C. Blue flowers develop their color based on the environment.
D. White flowers develop their color based on the environment.

Labels should be added to Figure 1 to clarify the color of each flower in the diagram because the colors cannot be determined in grayscale.

K. Preferred Editorial Styles

Spelling and Plurals

Common nouns
The plurals of most nouns are formed by adding *s*: boys, trees. Exceptions to this rule include the following:

- The plurals of words that end in ch, j, s, sh, ss, x, or z are formed by adding *es*: churches, biases, wishes, classes, foxes, waltzes.
- The plurals of words that end in a consonant and y are formed by changing *y* to *i* and adding *es*: babies, realities.
- The plurals of words that end in o are formed by adding es or s: heroes, potatoes, egos, cellos.
- The plurals of words that end in f or fe are usually formed by changing *f* to *v* and adding *es*: hooves, lives, but dwarfs, roofs.

Note: Many nouns have irregular plural forms (child/children, deer/deer, die/dice). When in doubt about the form or spelling of a plural, consult a dictionary.

Compound nouns
The plural of a hyphenated compound noun is usually formed by adding s to the main noun in the compound: brothers-in-law, courts-martial.

For solid, or closed, compound nouns, plurals are formed the regular way: classrooms, cupfuls, stopwatches.

The plurals of open compound nouns are formed by pluralizing the main noun: attorneys general, centers of industry.

Proper nouns
The plurals of proper nouns are usually formed by adding *s* or *es*. The plural of a proper noun ending in *y* takes an *s* (Monday/Mondays, the Smith family/the Smiths).

Letters, numbers, and abbreviations
The plurals of capital letters used as words, numerals used as nouns, and abbreviations are usually formed by adding *s*. To avoid confusion, the plurals of lowercase letters are formed by adding an apostrophe and an *s*.

Juan received all Bs on his report card.
the 1880s
DVDs
x's and *y*'s

Possessives

Singular nouns
The possessive of most singular nouns, both common and proper, is formed by adding an apostrophe and an *s*. This includes words that end with an unpronounced *s* and names with an ending pronounced *eez*.

the marquis's quarters
Albert Camus's novels
Euripides's works

However, a noun that is singular in meaning but plural in form takes an apostrophe only.

this species' characteristics
Hocking Hills' nicest campground

Plural nouns
The possessive of most plural nouns is formed by adding an apostrophe only.

the Joneses' house
the Martinezes' son
but
children's literature
women's rights

Letters and numbers
The possessive of letters and numbers is formed by adding an apostrophe and an *s*.

LBJ's diary
1980's worst flood

Joint vs. separated possession
When closely linked nouns are considered a single unit and "possess" the same thing, only the second noun takes an apostrophe and an *s*.

my mom and dad's house
Amelia and Brianne's teacher

When the things being possessed are not the same, both nouns take an apostrophe and an *s*.

my mom's and dad's birth certificates
Cleveland's and Chicago's rail systems

Compound Terms

compound noun: two or more nouns combined to form a single noun
compound modifier: a modifier that consists of two or more words

- An *open compound* is written as two words: real estate, sand dollar.
- In a *hyphenated compound*, the words are joined by a hyphen: self-esteem, half-baked.
- A *solid compound* is written as one word: playground, textbook.
- Avoid compound modifiers like "fish-eating bird." Instead use "a bird that eats fish."

Compound modifiers
Compound modifiers are usually hyphenated before a noun and open after a noun.

an open-ended question; a question that is open ended
a well-read student; a student who is well read
a 250-page book; a book that is 250 pages long
a sixteen-ounce bottle; a bottle that holds sixteen ounces

There are a few exceptions to this rule:

- When the compound modifier is a common open compound noun, it should be hyphenated only to prevent ambiguity.

  high school teacher
  real estate listing
  but
  short-story writer
  real-number theory

- When the first modifier in the compound is an adverb that ends with *–ly*, the compound is open.

  highly paid assistant
  hotly contested campaign

- When the compound is made up of a number and an abbreviated unit of measurement, the compound is open.

  a 5 km race
  a 3 m wall

- When a phrase is used as a modifier, it is usually hyphenated before a noun and open after a noun.

  over-the-counter medicine; medicine sold over the counter
  an up-to-date form; a form that is up to date

- When the second part of a compound modifier is omitted, a space follows the hyphen.

  fifteen- and twenty-year mortgages
  micro- and macro-evolution
  but
  third-, fourth-, and fifth-grade students

Prefixes and suffixes
Words that are formed with common prefixes and suffixes (*anti-, bi-, mid-, multi-, non-, over-, post-, pre-, re-, sub-, un-, under-, -fold, -less, and -like*) are usually closed.

  Bivalve
  Catlike
  Multipurpose
  Noninvasive

However, a hyphen should be used

- before a numeral or a capitalized word: post-1800, mid-September.
- before a compound term: non-self-sufficient.
- to separate combinations of letters that might be hard to read: anti-intellectual, de-ice, lava-like.

Use an en dash instead of a hyphen in a compound adjective when one of its elements consists of an open compound: post–World War II.

Capitalization
Proper nouns and adjectives are always capitalized.

Personal names and titles
All personal names (first, middle, last) are capitalized, as are initials, nicknames, and the suffixes Jr. and Sr. Do not set off a suffix with commas. Include a space between the initials in a personal name except when the initials are used alone.

Susan B. Anthony
Ivan the Terrible
E. B. White
Martin Luther King Jr.
LBJ

A person's title or office is capitalized only when it directly precedes a personal name and is part of the name.

President Lincoln; the president
Professor Johnson; the professor
Reverend Jackson; the reverend
General Grant; the general

When a title is used in apposition to a personal name (meaning it is used as a description rather than as part of the name), it is lowercase.

American president Lincoln
former president Bush
the Southern-born reverend Jackson

Kinship names
Kinship names are lowercase unless they directly precede or replace a personal name. When kinship names are used in apposition to personal names, they are lowercase.

My mom and dad have been married for 30 years.
Did you write to Aunt Kelly?
Can I have a cookie, Mom?
My kids love their aunt Kelly.

Racial and Ethnic names
Names of ethnic and racial groups are capitalized, as are adjectives derived from them. Do not hyphenate compound terms.

African Americans; African American poetry
Asians; Asian influence; an Asian American
Caucasians; Caucasian population

Geographic names
Proper names and nicknames are capitalized.

New York City
the Big Apple

Directional nouns are lowercase when they are used to indicate direction but capitalized when they refer to a distinct region.

a north wind; North African countries; in northern Africa
a southern climate; southern Ohio; the South; South America
eastern Illinois; the East Coast

Trademark and brand names
Use generic terms whenever possible. When using a brand name that is trademarked, capitalize the name but do not include the trademark symbol.

Post-it Notes; sticky notes
Kleenex; tissue

Titles of Works

Capitalization
Use headline style capitalization for titles of works. Capitalize the first and last words of the title and all interior words except

- articles (a, an, the).
- coordinate conjunctions (and, but, for, or, nor).
- prepositions, regardless of length, unless they are functioning as nouns, adjectives, or adverbs.
- the word *as.*
- scientific terms/names that begin with a lowercase letter (pH) or are lowercase in running text (*E. coli*).

Driving through Vermont
"The Ins and Outs of Trail Running"
"Reading for Fun"
Turn Up the Volume
A Primer on Soil pH

Hyphenated compounds in titles
Use the following guidelines for capitalizing a hyphenated compound in a title:

- Capitalize the first element of the hyphenated compound.
- Capitalize any subsequent elements unless they are articles, coordinating conjunctions (and, but, for, or, not), or prepositions.
- If the first element is a prefix or combining form that could not stand by itself as a word (anti, pre), do not capitalize the second element unless it is a proper noun or proper adjective.
- Capitalize the second number in a hyphenated number or fraction that is spelled out.

Heights of Sixth-Grade Students
Teacher-to-Teacher Initiatives
E-learning for Students
The Animals of Sub-Saharan Africa
Twenty-First-Century Skills
The Two-Thirds Majority

Tense
Present tense is the default tense. Some sentences will be in future and past tense.

Treatment of Terms

Note: See "Treatment of Numbers" in Part III for information about the preferred treatment of numbers.

Avoid the use of these words and phrases on science assessments:

- believe
- create
- prove (science does not work to prove hypotheses or theories true—this is a mischaracterization of how science works)
- truth (science is asymptotic to truth)
- above and below (referring to tables or graphics)
- of the following

Emphasis terms
Emphasis terms are boldface at all grade levels.

> Select **all** the questions you could ask to help solve the dilemma.
> Select **two** characteristics that can be hereditary.

However, avoid using qualifying terms like "best," "most likely," etc. In the rare instances when these terms are used, they should be boldface.

> Which is the **most likely** reason the population of deer decreased?

Letters as letters
Letters referred to as letters in text are italicized.

> Liam has 2 plant pots labeled with the letter *E*, 2 plant pots labeled with the letter *F*, and 2 plant pots labeled with the letter *G*.

Contractions
Contractions can be used in selections and other material from outside sources (e.g., stimuli). However, contractions are not used in items.

Options
The term "option" is never used in a prompt.

> Incorrect: "Select the option" or "Which option"
> Correct: "Select the [sentence, element, design, etc.]" or "Which [sentence, element, design, etc.]"

Abbreviations

> Note: Except for abbreviated units of measure and forms of address, abbreviations are rarely used in items. Abbreviations are used in graphics when space is an issue.

Acronyms and initialisms
acronym: an abbreviation based on the initial letters of a term and pronounced as a word (NASA, OPEC)
initialism: an abbreviation based on the initial letters of a term and pronounced by spelling out each letter (AARP, DNA)

Acronyms and initialisms are usually set in all capital letters without periods. When an acronym or initialism is preceded by an indefinite article, the choice of a or an is based on the pronunciation of the abbreviation.

an HMO
an AARP newsletter
a DNA sample
a NASA initiative
an OPEC worker

Unless an acronym or initialism is extremely well known (e.g., IRS, PTA, NATO), spell it out the first time it is used and enclose the abbreviated form in parentheses after the spelled-out term.

The grade-level expectations (GLEs) for science are listed below.
The GLEs for English Language Arts (ELA) are listed in the next section.

Latin abbreviations
Use Latin abbreviations only in parenthetical text. The abbreviations most commonly used are e.g. (for example), etc. (and so on), and i.e. (that is). In text, these abbreviations are set in regular type.

Unapproved resource materials (cell phones, dictionaries, etc.) are not allowed during test sessions. Reference books (i.e., dictionaries, thesauri) are not allowed during test sessions.

Taxonomic/Systematic conventions
Genus and species names should always be written in italics, including where they appear in figures and tables.

Tyrannosaurus
Tyrannosaurus rex
T. rex

Genus and species abbreviations are written in a combination of italicized and non-italicized text.

cf. *Tyrannosaurus*
Tyrannosaurus cf. *rex*
Tyrannosaurus sp.
Tyrannosaurus spp.
Tyrannosaurus insertae sedis

Names for other systematic groups should not be written in italics (e.g., Plantae, Animalia, Abelisauroidea, Dromaeosauridae)

Genotypes
Genotypes and single allele letters should be italicized.*

Gg
BB
dd

*except when text formatting is not available within an interaction type (e.g., ETC dropdown options and table input boxes)

Geographic abbreviations
The names of states are spelled out in running text. Abbreviations are used where a zip code follows or in other contexts in which abbreviations are appropriate (e.g., acknowledgments, graphics, tables, lists). In these cases, use the two-letter postal abbreviations without periods.

Spell out United States when it is used as a noun; either the abbreviation U.S. or US may be used as an adjective.

> The campus is in Westerville, Ohio.
> Please mail the documents to PO Box 121, Cloverdale, VT, 00111.
> the president of the United States (*not* the president of the US)
> the U.S. Treasury Department

Time
Use capital letters without periods in indicate eras. (Note that BC and BCE follow the date, while AD and CE precede the date.)

> 55 BC
> AD 1066

Months are spelled out in running text but may be abbreviated in graphics. Use the following abbreviations:

Jan.	May	Sept.
Feb.	June	Oct.
Mar.	July	Nov.
Apr.	Aug.	Dec.

Days of the week are spelled out in running text but may be abbreviated in graphics. Use the following abbreviations:

Sun.	Thurs.
Mon.	Fri.
Tues.	Sat.
Wed.	

Times of day: Lowercase letters followed by periods are used for a.m. and p.m.

Millions of years ago (or million years ago): mya

> 500 mya
> 1 mya

Part II: Specific Style by Interaction Type

A. Multipart Items

Many items within a cluster will have more than one interaction for the student to complete. The types of interactions can differ throughout the item or the item can consist of one type of interaction repeated.

- Each interaction heading, or "part heading," is title case and **bolded** before the stem of the interaction.
- There should be two hard returns between the part heading and the stem.

Part A

Select a testable, scientific question that can be answered by performing an experiment with the setup shown in Figure 1.

Note: The preamble sentences "The following question has two parts. First, answer part A. Then, answer part B." are no longer used for our science items.

For the stem of subsequent interactions, do not refer back to the student's response to previous interactions ("… you chose in part A") unless there is dependent scoring. This referencing suggests that the student can get credit for pairing parts A and B correctly without selecting the correct answer to the question for part A.

Use:
Which given evidence supports the answer to part A?
OR
Which statement provides a reason for the answer to part A?

B. Multiple Selection (Multi-Select) Interactions

Multi-select items are selected response items that allow the student to choose more than one option.

Note: For multiple choice interactions, see "Selected-Response Interactions" in Part I.

Task Directions

- For all elementary and middle school interactions use "Select (number of options) …" The number should be spelled out and **bolded**.

  Select **four** statements that describe predator-prey relationships.

- For some high school interactions, the exact number of selections may be omitted.

  Select **all** the statements that describe predator-prey relationships.

Options

- If options are one word or phrase, use lowercase; no punctuation is necessary.
- Option order:
  - In the order they appear in a stem graphic or table
  - In ascending order, if numeric options
  - Alphabetically or ascending or descending order, if all options are one word
  - By ascending or descending order, if phrases
  - If graphics only, arrange in a logical order

Human Readable Rubric

The human readable rubric outlines correct responses to the item.

Example:

A full-credit (1 point) response includes
The student selected:
- "the flower is red"

AND
- "the flower is blue"

C. Edit Task Inline Choice (ETC) Interactions

ETC items allow students to select a response from a dropdown menu in order to complete a sentence or a cell in a table.

Task Directions
For one dropdown box, use: "Click on the blank box and select the word/phrase …"
For two or more dropdown boxes, use: "Click on each blank box and select a word/phrase …"
OR
"Click on the first blank box and select the word/phrase …. Then, click on the second blank box and select the word/phrase …."

> Note: When the student is directed to complete a model, causal chain, steps in a process, or similar task, do not use the phrases "in order" or "in sequence" as that is implied in the directive (the student would not be completing a chain of causality if the steps were out of sequence). (But see note in section H, External Copy Interactions.)

Options
The number of options in each dropdown will vary from dropdown to dropdown and from item to item based on content.

Dropdown options that are full sentences should have initial caps and end punctuation. In tables and other graphics, dropdown options that are fragments should have initial caps and no end punctuation.

The options within the dropdowns should be ordered using the same guidelines as multiple choice and multi-select interactions, most commonly in ascending length.

Human Readable Rubric
Example:

A full-credit (1 point) response includes
The student selected:

- "increased" for the first blank

AND

- "increased" for the second blank

OR

- "decreased" for the first blank

AND

- "decreased" for the second blank

Exemplar: Include the correctly completed sentence, diagram, or table.

D. Table Interactions

Table Match Interactions

Table match interactions allow the student to select cells within a table to show a relationship between the column header and the row header.

Task Directions

- Use: "Select the boxes to identify each organism's role in the ecosystem." OR "Select the boxes to show the order of the steps of (process) …"
- When there is more than one correct answer combination in an item, the following guideline **may** be included:
    - There may be more than one correct answer.
- When it is unclear that a student can and should select more than one box, the following guideline should be included:
    - You may select more than one box for each part …

Tables

- Column and row headings may contain text or graphics. When graphics are present, it is preferable for a text label or descriptor to be present.
- In the directions, refer to the first column first (when possible).
- If the text in the column or row heading is a complete sentence, it should begin with a capital letter and end with the proper punctuation.
- If a column or row heading is a phrase or one word, it should begin with a capital letter and should not have punctuation.
    - The subsequent words of a phrase should not be capitalized, unless the phrase is a proper noun or adjective.
- Within an item, column headings must all follow either title case/capitalization or sentence case. Within an item, row headings must all follow either title case/capitalization or sentence case.
- Title and sentence case/capitalization should be used consistently. Column and row headings should be in one of the following orders when they appear:
    - In the order they appear in a stem graphic or stem table
    - In ascending or descending order if all options are one word
    - By length (ascending/descending), if phrases
    - If graphics only, arrange in logical order
- Column headings should be bolded, centered, and boxed and have no shading.
- Row headings should be normal weight text and have no shading.
    - Left align row headings when one or more headings have more than one word
    - Center row headings that are a single word or number

Human Readable Rubric

- Use quotes to reference the row and column headings in the interaction.
- Row and column heading references are capitalized as they are in the interaction.
    - Row and column headings with no text should be described. Example: the picture of a new moon
- Example:

  A full-credit (1 point) response includes
  The student selected:
    - "Blue" and "Green" for "Reflected"

  AND
    - "Orange" and "Yellow" for "Absorbed"

  AND
    - Nothing else

Exemplar: Include a correctly completed table.

Table Input Interaction
Table input interactions allow the student to enter numbers, symbols, and/or words into the cells of a table.

Task Directions

- Use: "Enter the [number, value, etc.] in the blank box."
- Let the student know if there is more than one correct answer.

Note: For table specifications, see the "Table" section of "Table Match Interaction."

Human Readable Rubric
Example:

A full-credit (1 point) response includes
The student entered:

- "8" for column 2, row 2

AND

- "12" for column 3 row 3

Exemplar: Include a correctly completed table.

E. Hot Text Interactions

Hot text interactions allow the students to click on text to “select” it.

Task Directions

- Use “Select the [words, sentences, etc.] that …”

Human Readable Rubric
Example:

A full-credit (1 point) response includes
The student selected:
- “nucleus”

AND
- “cell wall”

F. Graphic Response Interactions

Graphic response or "grid" interactions allow students to drag/ drop, draw dots or lines, and/ or use hot spots to complete a diagram or model.

Note: See the "Graphs" section of Part IV for graph specifications.

Task Directions

- Response boxes contained in the grid background are referred to as "blank boxes" in task directions and guidelines.
- References to "correct," "appropriate boxes," etc., in the task directions and guidelines should be avoided.
- To help the student navigate the various components of the item, use consistent language in the stem, palette objects, and answer space.
  - For example, if the item is about Jupiter, the various components of the item should use the word "Jupiter" and not the more general reference to "planet."
  - If the palette objects are referred to as "birds" in the task directions, they should be referred to as "birds" in the guidelines as well.
  - This consistency should be extended to the scoring assertion rationales.
- Whenever possible, refer to the material in the answer space in specific terms ("in the picture," "in the table," "in the graph," etc.). Do not use the terms "graphic," "image," or "answer space."
- References in the stem to diagram headings and grid background titles should be in the same capitalization as in the headings and titles but should not be bolded.
- Whenever possible, refer to the palette objects and groups of palette objects in specific terms (e.g., "chemical elements").
  - Reference to two or more palette objects should be lowercased.
  - If the palette objects are just pictures, lowercase the references to them in the stem.
  - Reference to a specific palette object in the stem should use the same capitalization as in the palette object labels.
- If it is not possible to refer to palette objects in specific terms, then:
  - A palette object that includes a picture, with or without a label, is referred to in the item as an "object/picture/diagram."
  - A palette object that is text only can be referred to as a "label/name."
  - A mixed collection of palette objects can be referred to as "object/label."
  - **Do not** use the term "palette object."
- **Palette Bar Drag and Drop Interactions:** For interactions where palette objects appear in the left palette bar, use the verb and preposition "Place … in …" in task directions. Also, refer to boxes with dashed lines as "blank boxes." (Note: Do not use the term "drag.")
- **Pre-placed Drag and Drop Interactions:** For interactions where palette objects are preset in the gray pre-placed box in the grid background, use the verb and preposition "Move … into …" in task directions. Also, refer to boxes with dashed lines as "blank boxes." (Note: Do not use the term "drag.")
  - Pre-placed interactions should **not** have a Delete tool. The student will not be able to get the palette object back once it is deleted, so the tool is not needed.
- **Hotspot Interactions:** Interactions that allow students to select multiple words, text graphics, etc. on the background in order to complete their machine scored constructed response. For these interactions, use the verb "Select …" in task directions.
- **Hotspot Bar Graph Interactions:** Interactions that allow students to construct a bar graph. For these interactions, use "Click on a line to show where the top of each bar should be" in the task directions.
  - Note: Bar graph interactions should have only 'select' hotspots. Do not add 'hover' hot spots for any grade level. This is consistent with AIR math item development.
- **Graphing Interactions:** Interactions that require the student to use the "Add Point," "Connect Line," and/or "Add Arrow" tools. For these interactions, use the verb "Use …" in

the task directions; also, use the term "tool," not "button" (e.g., "Use the Add Point tool to …").

    - Note: Interactions that contain the "Connect Line" and/or "Add Arrow" buttons must also have the "Add Point" button. Buttons should be in the following order: Delete, Add Point, Connect Line, Add Arrow (single), Add Arrow (double).

- When applicable, the following guidelines should be included:
    - "Use only **one** [palette object] in each blank box you fill in."
        - Note: this is used in all drag and drop items in which only one answer is required or fits in each blank box. Also, the word "one" should be bolded.
    - "The [palette objects] may be used once, more than once, or not at all."
    - "The [palette objects] may be used more than once."
    - "Not all [palette objects] may be used."
        - Note: this is used when the number of palette objects equals or exceeds the number of blank boxes.
- The guidelines above may vary in language in order to reference specific aspects of an interaction.
    - Use specific references to the object, rather than the word "object," when possible. If not possible, use "object."
    - For items that have only labels and no objects, use "label."
    - For items that have only one box to fill in, use "the blank box."

Human Readable Rubric

- It is not necessary to use quotes for references to palette objects or graphic labels in written rubrics.
- Palette objects with text are capitalized as their labels are capitalized, and are not described as "the 'Maple Tree' object."
- Palette objects or response box locations with no text should be described. Example: "the picture of a new moon."
- Example:

    A full-credit (2 point) response includes
    The student placed:
    - Only the red bird in the smallest birdhouse region

    AND
    - Only the black bird on the power lines OR on the roof

- Exemplar: Include a snapshot of the correct response.

G. Simulation Interactions

Simulation interactions allow the student to investigate a phenomenon by selecting variables to get output data. Some simulations are accompanied by animations.

Task Directions
Simulation directions should be as clear and simple as possible.

- State the task clearly and concisely.
- Use either "controlled experiment" or "investigation" when discussing the type of activity the student is using the simulation for.
- Include the directive to either "Click on Start to [run a trial, see the results, etc.]" or "Click on Run to [run a trial, see the results, etc.]" depending on whether the button says "Start" or "Run."
- State the number of trials that can be run.
- A guideline should always be included informing the student as to whether they can delete trials using the trash can icon or not.
- Provide the student enough direction that they will provide the information needed for scoring.

Examples (prompt only):

Use the simulation to gather measurements that can be used to predict the weather.

Use the simulation to test prototypes for the design.

Example (prompt with guidelines):

Use the simulation to conduct a controlled experiment to determine what makes the balloon float.

- You will be limited to **five** trials.
- Click on Start to run a trial.
- Click on the trash can at the end of the row to delete a trial and generate new data.
- You will be scored only on the trials present in the simulation table.

Layout
Layouts of the simulation will vary due to content being investigated.

Animations
Animations within a simulation should follow the standard animation guidelines. Animations should be as short as possible in order to decrease the amount of time spent on the simulation by the student. Use the directive to "Click on the small gray arrow to start the animation …" OR "Click on the small gray arrow to watch [Earth revolve around the sun, the rabbit come out of its hole, etc.]."

- Animations within a stimulus should be no wider than 350 pixels to avoid horizontal scrolling.

Output Table
Output tables should follow the same format and style of data tables presented in text.

Human Readable Rubric

- Students are scored on their decisions during the investigation.
- The combination of variables needed for each score point should be listed.
- Example:

A full-credit (3 point) response includes

The student ran trials with the following variables:

- "blue," "heavy," and "square"

AND

- "yellow," "heavy," and "square"

AND

- "blue," "light," and "square."

- Exemplar: Include a snapshot of a correct output data table.

H. External Copy Interactions

External copy items allow the student to select text from the stimulus to copy into the answer space.

Task Directions

- Click on the first pencil icon.
- Then, click on a highlighted step [or "section from the passage"*] to make your first selection. Click on the other pencil icon(s) to make your remaining selection(s).
- Click on the circular arrow that follows any selection you would like to change.

**Or "source(s)," "part C," or any other text that will become highlighted by the pencil; do not use "stimulus."*

> Note: When the student is directed to complete a model, causal chain, steps in a process, or similar task, and the steps are given in a list, cite in the introduction to the list and in the item prompt that they are out of order and should be placed in order. (But see note in section C, Edit Task Inline Choice [ETC] Interactions.)

Human Readable Rubric

The human readable rubric outlines correct responses to the item.

Example:

> A full-credit (1 point) response includes
> The student selected:
> - "the flower is red"
>
> AND
> - "the flower is blue"

E. Equation (EQ) Editor Interactions

Equation (EQ) Editor interactions allow the student to enter numbers, symbols, etc. into the answer space by either typing them manually or by using the built-in keyboard.

Task Directions

- Use "Enter the [equation, number, value, etc.] in the blank box."

Human Readable Rubric
Example:

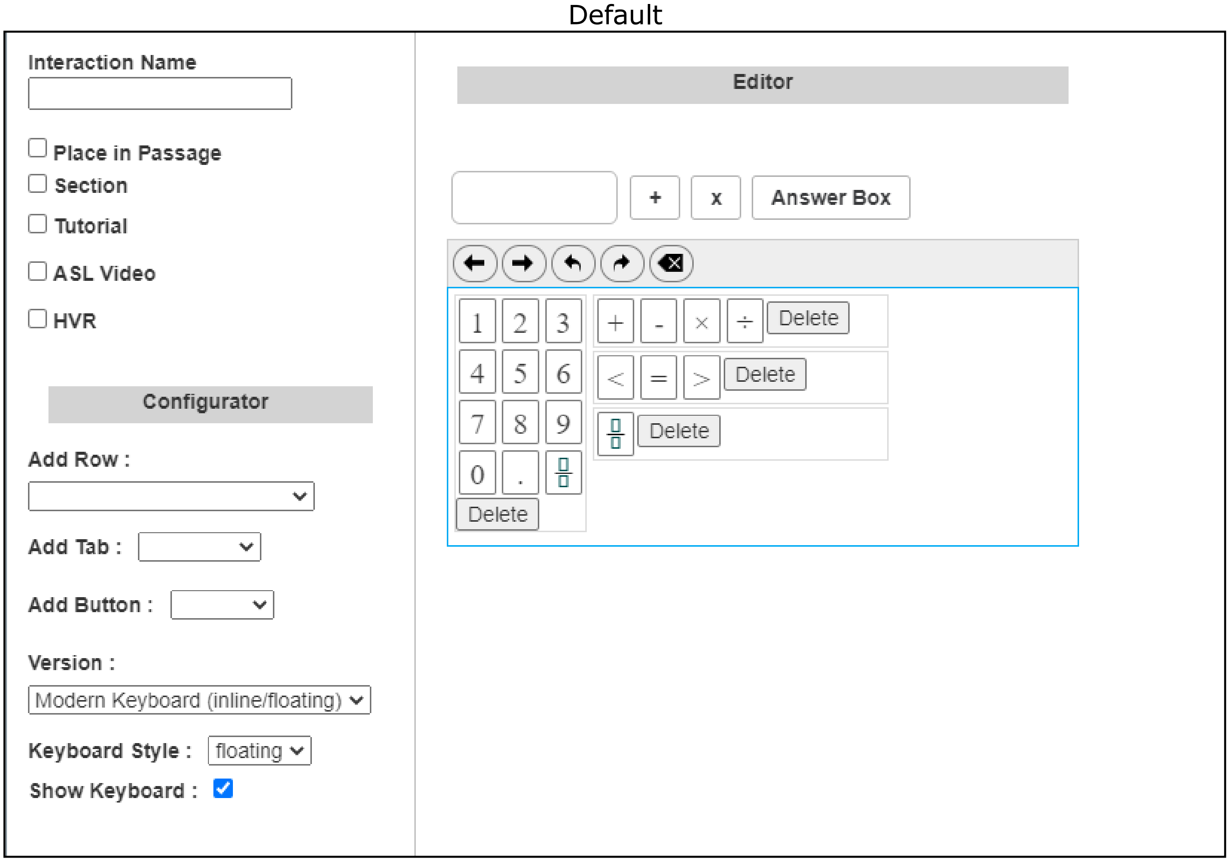
A full-credit (1 point) response includes
The student entered the equivalent of:
- "10.512"

AND
- "5*k* – 20*y*"

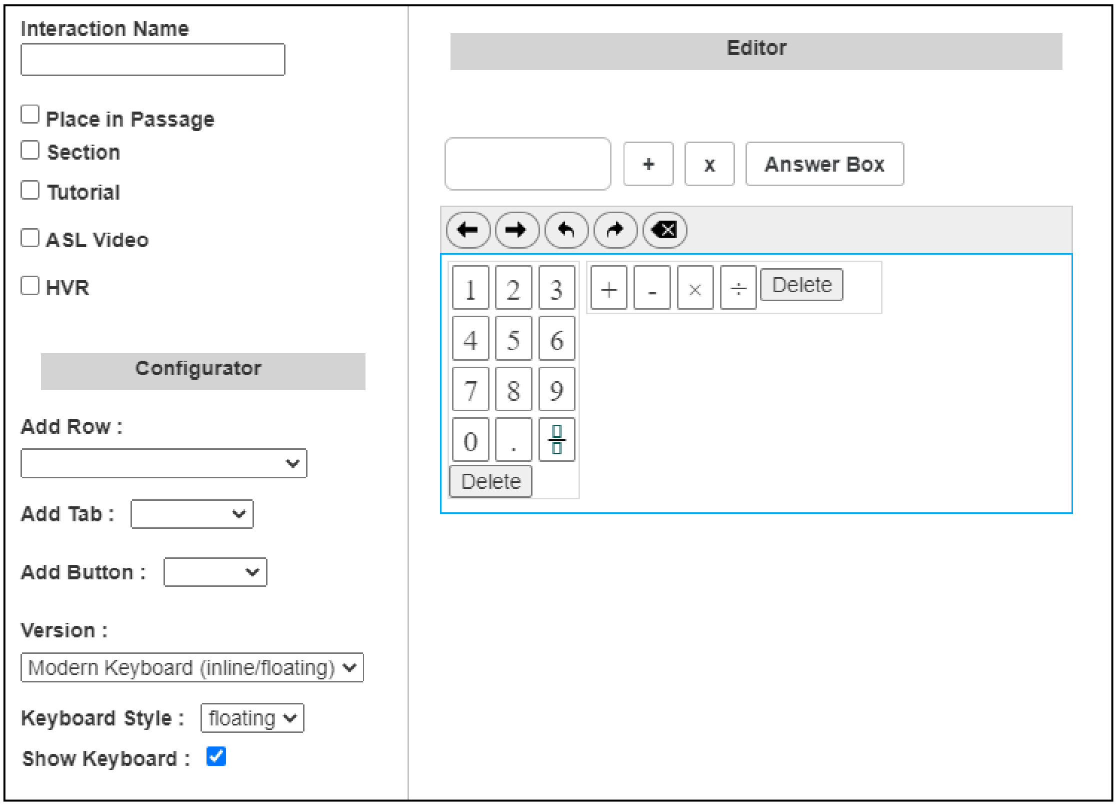
Keyboard
The keyboard is tailored to what the student needs to complete each interaction. Always include the absolute minimum number of buttons that the student will need.

The default keyboard includes a number pad, two fraction buttons, operators, and inequality symbols (see Default graphic below). *Always remove the standalone fraction button.* Also, remove the inequality symbols if they will not be used. This leaves the standard keyboard with the number pad, one fraction button, and the operators (see Standard—Revised Default graphic below).

Default



Standard—Revised Default

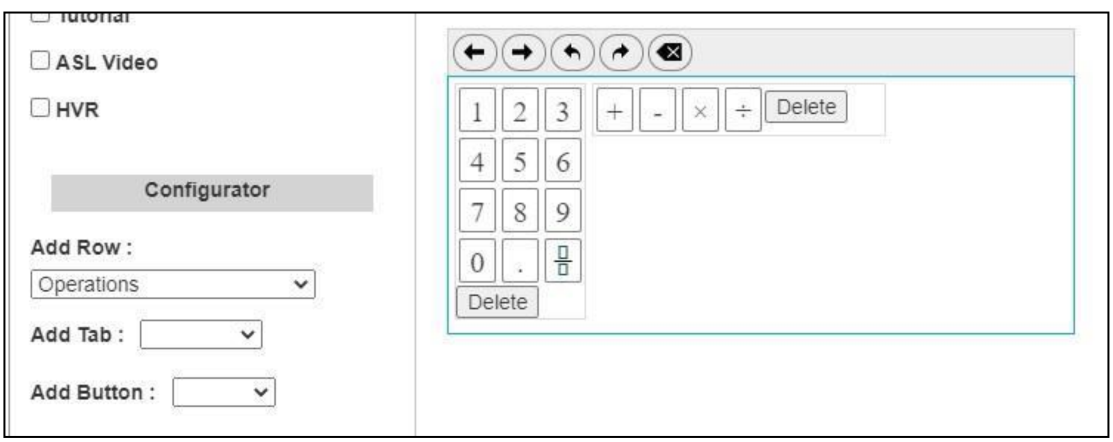


Operators at Different Grade Levels

- To change the operators on the keyboard, click “Add Row.”

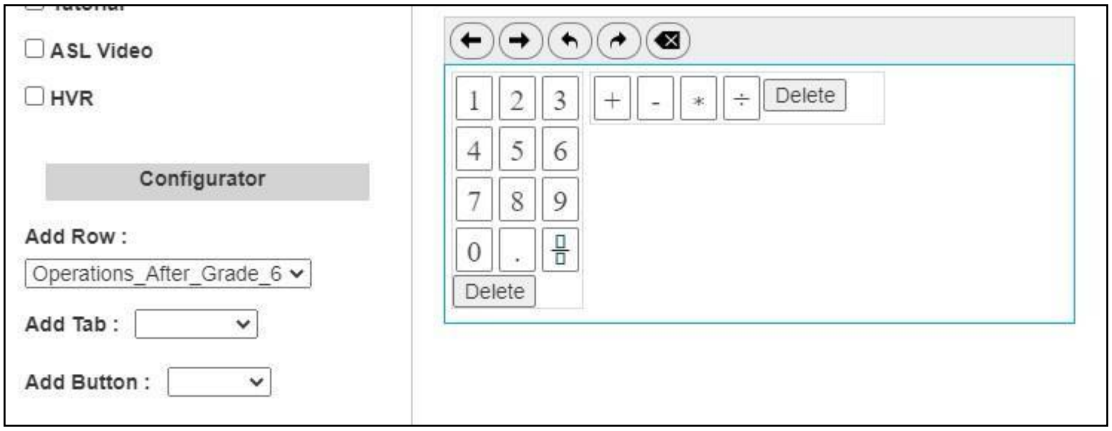
Operators for grades K–5

- This is the default. It can be selected as “Operations.”
- The multiplication symbol is x.



Operators for grades 6–12

- Select "Operations_After_Grade_6."
- The multiplication symbol is *.



F. Data Exploration Interactions (TUVA Graphs)

Task Directions

Graph 1 allows you to select which data to place along the *x*- and *y*-axes. You may make several graphs as part of your investigation of the data.

- To make a graph, click on an attribute and then click in the blank box that represents either the *x*- or *y*-axis.
- You can change the way the graphed data appear by using the toolbar at the top.
- You may place more than one attribute on each axis by dragging the attribute to the small "+" sign to the left of or below the current attribute.
- To change your selection, click on the small "x" to remove an attribute from that axis, and then add a new one.

Part III: Number Treatment Style Conventions

A. Treatment of Numbers

Note: This section provides GENERAL guidelines for the treatment of numbers. These guidelines are applicable to ALL content areas.

Words vs. Numerals
Use **words** for

- Numbers zero through nine, with the exceptions in this section.
- Numbers that appear as the first word in a sentence (content specialist determines exceptions).

Acceptable
10 mice are white.

Preferred
Ten mice are white.
OR
There are 10 white mice.

Use **numerals** for

- numbers 10 and above.
- numbers that precede abbreviated units of measure.
- numbers that precede or follow symbols: 10%, $20.00.
- numbers that appear in equations/expressions.
- numbers used to solve mathematical problems.
- numbers included in parts of published works: volume 2, chapter 4.
- dates and years: 1000 BC; December 1, 1975.
- times of day that precede the abbreviations a.m. and p.m.: 11 a.m., 3:00 p.m.
- telephone numbers.

In lists and series of numbers, use either words or numerals consistently.

Pat observes the following insects: 10 ants, 5 flies, and 2 crickets.
A circle graph is divided into five sections labeled 3, 6, 9, 12, and 15.

Note: The content specialist determines whether to use words or numerals when guidelines for the treatment of numbers contradict one another.

Ordinals

- Spell out ordinals first through ninth.
- Use numerals for ordinals 10th and above; when numerals are used, suffixes are set on the baseline, not in superscript.
- In lists and series of ordinals, use either words or numerals consistently.

The student finished the race in second place.
The 25th customer to enter the store today will win a prize.

Commas in Numbers
Use a comma in

- numbers with five or more digits: 50,000.
- numbers with four digits only if other numbers in the item have five or more digits.

The teacher has driven his car a total of 28,000 kilometers in three years. He drove 6,000 kilometers the first year.

- numbers written as words: one million, eighty-seven thousand, three hundred twenty-two.

Do **not** use a comma in

- numbers with four digits (unless other numbers in the item have five or more digits).

  The teacher drove her car 836 kilometers in August, 1027 kilometers in September, and 914 kilometers in October.

- compound measures, such as height measurements: 12 meters 50 centimeters—not 12 meters, 50 centimeters. (Note: In most cases, use 12.5 meters.)

Values Less Than One

- Use singular units of measure with values less than one: 0.25 gram, not 0.25 grams.
- Include a zero before the decimal point in decimal values less than one: 0.15, not .15.

Negative Numbers

- Use a mid-point en dash to indicate negative numbers. Which point has the coordinates (2, –5)?

Fractions

The content specialist determines whether fractions are spelled out or expressed as numerals.

- As words, fractions are hyphenated as nouns, adjectives, and adverbs. Two-thirds of the students in the class ride the bus.
- As numerals, fractions are stacked vertically and appear at 90% of the base text size:

14 pt. = 12.6 pt.

A student gave $\frac{1}{4}$ of her sandwich to her friend.

Percentages

- Use the word percent after a number word. Five percent of the dogs have spots.
- Use the word percentage, not percent, as a stand-alone term. What percentage of the cats are white?
- Use the percent symbol after a numeral: 5%. (See "Symbols and Special Characters" in Part V for the preferred styles for symbols and special characters.)

  Of the marbles in the jar, 40% are red and 60% are blue.

Exponents

Exponents and other superscripted characters are scaled to 70% of the base text size: 14 pt. = 9.8pt.

$s^2 \times 7 = 28$

Ratios

Use a colon in ratios. Do **not** insert a space before *or* after the colon.

The ratio of solute to red solvent is 1:4.

Coordinates and Ordered Pairs

- Enclose coordinates and ordered pairs in parentheses.
- Include a comma, followed by a space, after the first number.

  Point A has the coordinates (3, 4).

- Include a space after the name of a point that precedes coordinates or ordered pairs.

  Line m begins at point A (2, 5) and ends at point B (−1, −3).

Dates

- *Abbreviated years:* Avoid abbreviating years whenever possible. When a year is abbreviated, the first two numbers are replaced by an apostrophe (not an opening single quotation mark): the blizzard of ’76.
- *Months and days:* In running text, dates are written in the following form: February 10, 2012.
- *Centuries:* Centuries are spelled out and lowercase: the twenty-first century, nineteenth-century literature.
- *Decades:* Decades can be spelled out or expressed as numerals; if spelled out, they are lowercase: the nineties, the 1990s. (Note that no apostrophe is used in the plural form of decades.)
- *Eras:* Eras are expressed as numerals: 55 BC, AD 1066. (Note that BC and BCE follow the date, while AD and CE precede the date. All four abbreviations are uppercase with no periods.)

Times of Day

The content specialist determines how to present times of day in individual test items. The following conventions should be applied based on the presentation selected:

- Use numerals with the abbreviations a.m. and p.m. (Note that the abbreviations are lowercase with periods.) It is redundant to include phrases such as “in the morning,” “in the afternoon,” or “at night” after a.m. or p.m.

  The student wants to see a movie that starts at 4:10 p.m.

- Spell out numbers used with the term “o’clock.”

  The student leaves for school at eight o’clock.

- To avoid confusion, spell out the terms noon and midnight (in place of 12 a.m. or 12 p.m.).

  The student works from 6:30 p.m. to midnight.

B. Equations / Expressions and Patterns

General Guidelines

- In general, equations and patterns are 14 pt. Verdana; however, the font sometimes varies for equations and patterns that include symbols.
- In introductory statements, use the term "equation" or "expression" to refer to an equation or expression. Do not use the term "number sentence."
- Use the term "pattern" to refer to patterns of numbers and patterns of symbols.
- Use the term "step" to refer to the position of a term in a pattern: the fifth step in the pattern.
- In items, equations are center-aligned and can be given the headings "Equation 1," "Equation 2," etc.
- See "Graphics and Other Stimuli in Items" in Part I for additional guidelines.

Variables and Symbols

- In general, variables are lowercase and italicized. (However, variables in provided formulas can be uppercase or lowercase, as tradition and context dictate.)

  Solve for *x*.
  formula for area: $A = lw$

- In **grades 3–5**, use boxes to indicate missing/unknown values in equations.
- In **grades 6 and above**, use variables or boxes to indicate missing/unknown values in equations.

  $6 + n = 12$

- In all grades, use a question mark or underscored blank space to indicate missing terms in patterns. (The content specialist determines whether to underscore the question mark.)

  2, 4, ?, 8, 10

Operational Symbols

See "Symbols and Special Characters" in Part IV for a complete list of operational symbols used in Mathematics as well as the preferred styles for symbols and special characters used in item text and graphics. See "Words vs. Symbols" in this section for information about using words and symbols to identify geometric objects in running text.

Multiplication symbols.

- In grades 3–5, use the multiplication symbol.

  8×7

- In grades 6 and above, use the product dot or do not include a symbol. (Do not use the × symbol, except in scientific notation, to avoid confusing with the variable *x*.).

  $8 \bullet 7$
  $(10 - 2)(7)$

- In all grades, use the multiplication symbol in scientific notation.

  5.02×10^6

C. Units of Measure

When to Abbreviate

- Spell out units on their first usage in a cluster stimulus or in an item stem followed by the abbreviation in parentheses.
- Abbreviate units thereafter in item stems or options.

Stephanie has 15 pieces of string. Each piece is 5 meters (m) long. How many meters of string does Stephanie have altogether?
A. 25 m
B. 50 m
C. 75 m
D. 100 m

- Spell out units in tables (See "Units of Measure" in Part IV for additional information.)
- Abbreviate units in graphics. (See "Units of Measure" in Part IV for additional information.)

Abbreviations

Note: Do not include periods in abbreviated units of measure.

Metric Units

Table 4 shows the correct abbreviations for metric units of measure.

Table 4. Abbreviations for metric units of measure

Unit	Abbreviation
Millimeter	mm
Centimeter	cm
Meter	m
Kilometer	km
Milligram	mg
Gram	g
Kilogram	kg
Milliliter	mL
Liter	L

kilogram-meter per second = kg·m/s

Here is a link to SI units to use as a resource:
https://www.nist.gov/pml/weights-and-measures/metric-si/si-units

Hyphenated abbreviations used with numbers as modifiers are not hyphenated.

Example: a 10 m pole, a 10-meter pole

Temperature Units

Table 5 shows the correct abbreviations for units of temperature.

Table 5. Abbreviations for units that measure temperature

Unit	Abbreviation
degrees Celsius	°C
Kelvin	K

Unit	Abbreviation
degrees Fahrenheit	°F

Note: Degrees Fahrenheit is preferred when discussing weather or body temperature.

Time Units

Table 6 shows the correct abbreviations for units of time.

Table 6. Abbreviations for units that measure time

Unit	Abbreviation
Day	day
Minute	min
Second	s

Unit	Abbreviation
Hour	hr
Month	mo
Year	yr

Plural Units

- Do not add an "s" for plurals of abbreviated units.

  Melissa is making 10 identical saltwater solutions. She needs 100 milliliters (mL) of distilled water to make 1 solution. How many milliliters of distilled water does she need to make 10 solutions?
  A. 10,000 mL
  B. 1,000 mL
  C. 100 mL
  D. 10 mL

- Use a singular verb with physical quantities. How many grams (g) of silver is produced?

Punctuation and Spacing

- Do not include periods in abbreviated units: cm, s, cm, m/s^2
- Do not include commas in compound measures, such as height measurements: 12 meters 50 centimeters—not 12 meters, 50 centimeters. (Note: In most cases, use 12.5 meters.)
- Include a space between numerals and abbreviated units, except in temperatures: 30 cm, but 90°F.
- In temperatures, do **not**
  - include a space between the numeral and degree symbol, or between the degree symbol and the unit: 0°C.
  - use a degree symbol with the abbreviation for kelvin: 223K not 223°K.
  - There is also NO space between number and degree symbol or directional in longitude or latitude measurements: 90°N.

Square and Cubic Units

- When units are spelled out, spell out the terms square and cubic.
- When metric units are abbreviated, use superscript to show square and cubic units.

  The student drew a square with a side length of 6 centimeters (cm). What is the area, in square centimeters (cm^2), of the student's square?

  A. 15 cm^2

  B. 30 cm^2

  C. 36 cm^2

  D. 54 cm^2

- When customary units are abbreviated, use sq and cu for units.

  83 sq m

Conversions

The content specialist determines whether to include conversions in items.

- When included, conversions are enclosed in brackets after the punctuation mark at the end of the stem. The conversion itself includes no punctuation.

  What is the volume of the rectangular prism?
  [1000 milliliters (mL) = 1 liter (L)]

- Use an equal sign in conversions that involve units of measure; always position the value with the smallest unit on the left side of the equal sign.

  What is the area, in square meters (sq m)?
  [100 centimeters (cm) = 1 meter (m)]

- Use the term "represents" in conversions that involve scales or that assign a value to a graphic.

  What is the total area of the grid? [" represents 1 unit]

Pi

If an approximation for pi is desired to be given in a particular item, the value of pi should also be enclosed in brackets at the end of the item stem. However, the value is stated as a sentence that ends with a period.

What is the height of the cylinder? [Use 3.14 for π.]

D. Preferred Language

Conditional ("if") Clauses
Recast conditional clauses (e.g., "If this happens …?") when possible. If a conditional clause cannot be avoided, position it at the end of the sentence.

> Let x = 7. What is the value of y?
> *not*
> If x = 7, what is the value of y?
>
> The pattern continues. Which shape will be in step 25?
> *not*
> If the pattern continues, which shape will be in step 25?

Table vs. Chart

- Use "table" when data are organized and related in some way.

  The student recorded the measurements in the Volume of Water vs. Height of Plants table.

- Use "chart" when data are not organized to emphasize comparison among discrete items or related in any way (e.g., data that are listed).

  The Fancy Pens chart lists the types and colors of pens the student can buy.

Percent vs. Percentage vs. %

- Use the word "percent" after a number word.
- Use the word "percentage" as a stand-alone term: a percentage of students.
- Use the percent symbol after a numeral.

  A student has 20 folders. Each folder is either red, green, or yellow. Twenty percent of the folders are red. Forty percent of the folders are green. What percentage of the folders are yellow?
  A. 20%
  B. 30%
  C. 40%
  D. 50%

Constructed-Response Items

- Include units of measure in the stem so that students are not penalized for omitting units from their responses. Also, CR stems should be worded as imperatives (e.g., "Calculate the area"), not questions (e.g., "What is the area?").

  Calculate the area, in square meters (sq m), of the garden.

- In sentences that refer to item parts, the word "part" is lowercase and the part letter is capitalized.

  Identify the statement that supports the choice in part A.

- The content specialist determines the language of statements that ask students to justify their answers.

  Show your work. Explain your reasoning.
  Show or explain how you got your answer.
  Show or explain how you know your answer is correct.

Part IV: Graphic Specifications

Note: The specifications in this section are for all graphics, including graphics used in technology-enhanced interactions (TEIs).

A. Text Elements

This section provides specifications for text elements that appear in graphics.

Fonts

In general, text in graphics is Verdana. However, exceptions are made for graphics that require a special look (e.g., advertisements, posters). Table 7 shows general font specifications for different text elements in graphics.

Table 7. General font specifications for text elements in graphics

Font Specifications	
Text Element	**Font**
Title	- 14 pt. Verdana Bold - Title case
Headings (e.g., axes headings, column headings)	- 14 pt. Verdana Bold - Title case
Labels and text	- 14 pt. Verdana - Sentence case
Credit lines	- 10 pt. Verdana - Lowercase

Note: At this time, Verdana is specified as the primary font for test content.
However, another font may be chosen upon further analysis of the effects that fonts have on readability and students' ability to retain information.

Symbols and Special Characters

The table that follows shows the preferred styles for symbols and special characters. These specifications apply both to symbols in graphics and symbols in text, with the exceptions noted.

- In graphics, the size and style (e.g., boldface, italic) of a symbol depends on where it appears in the graphic. For example, a symbol that is part of a title is 14 pt. and boldface. Use Table 7 in the previous section to determine the correct size and style of symbols.
- In items, symbols are the same size and style as the surrounding text.

Table 8. General font specifications for symbols and special characters in graphics and text

Font Specifications for Symbols and Special Characters		
Symbol/Character	**Font**	**Description**
&	Verdana	- Ampersand - Do not use in science
©	Verdana	- Copyright symbol - Used in acknowledgments and credit lines
$2.00 50¢	Verdana	- Dollar sign/cent symbol - Used in dollar amounts
%	Verdana	- Percent symbol - Used in percentages
’	Verdana	- Smart (curly) apostrophe - Used in possessives
′	Verdana	- Prime mark - Used to indicate prime numbers
‘	Verdana	- Okina - Glottal stop used to spell Hawai‘i
0°C 45° angle	Verdana	- Degree symbol - Used in temperatures and angle measures - No space between number and degree symbol or between degree symbol and unit of measure
+	Verdana	- Addition symbol - Used in equations/expressions
–	Verdana	- En dash - Used as subtraction symbol in equations/expressions; also used in number ranges and with negative numbers
×, •	Verdana	- Multiplication symbol and product dot - Used in equations/expressions

Table 8. General font specifications for symbols and special characters in graphics and text *(cont.)*

Font Specifications for Symbols and Special Characters		
Symbol/Character	**Font**	**Description**
÷	Verdana	- Division symbol - Used in equations/expressions
=	Verdana	- Equal sign - Used in equations/expressions
$\frac{1}{2}$	Verdana	- Vertically stacked fraction - Scaled to 90% of text size: 16 pt. = 14.4 pt.; 14 pt. = 12.6 pt.
π	Symbol Std.	- Pi - Not italicized - Used in equations/expressions
~	Verdana	- “Similar to” symbol - Used to indicate similar lines, shapes, and angles
(4, 3)	Verdana	- Coordinates and ordered pairs - Enclosed in parentheses - Comma, followed by space, after first number
1:2	Verdana	- Ratio - No space before *or* after colon
$V = l \times w \times h$	Verdana	- Variables - Uppercase or lowercase, as tradition and context dictate - Italicized - Used in equations/expressions and formulas
x, y	Verdana	- x-axis and y-axis labels - Lowercase - Italicized - Used to label x- and y-axes in line graphs, scatter plots, and coordinate grids

Table 8. General font specifications for symbols and special characters in graphics and text *(cont.)*

Font Specifications for Symbols and Special Characters		
Symbol/Character	**Font**	**Description**
A, B, C	Verdana	- Point label - Boldface, italicized, uppercase letter (in graphics only; see "Points" in Part IV for point labels in text) - Used to label points and other geometric objects
1st, 2nd	Verdana	- Ordinals - Positioned on baseline (not superscripted)
13^2	Verdana	- Superscript - Scaled to 70% of text size: 16 pt. = 11.2 pt.; 14 pt. = 9.8 pt. - Raised by 33% with a baseline shift of +6
H_2O	Verdana	- Subscript - Scaled to 70% of text size: 16 pt. = 11.2 pt.; 14 pt. = 9.8 pt. - Lowered by 33% with a baseline shift of –6
9:00 a.m. 3:00 p.m.	Verdana	- Used to indicate times of day - Lowercase (not small caps)
100 BC/BCE AD/CE 1800	Verdana	- Used to indicate eras, epochs, etc. - Uppercase (not small caps)

Note: Symbols and special characters are used at the content specialist's discretion.

B. Graphic Size

Graphics should be

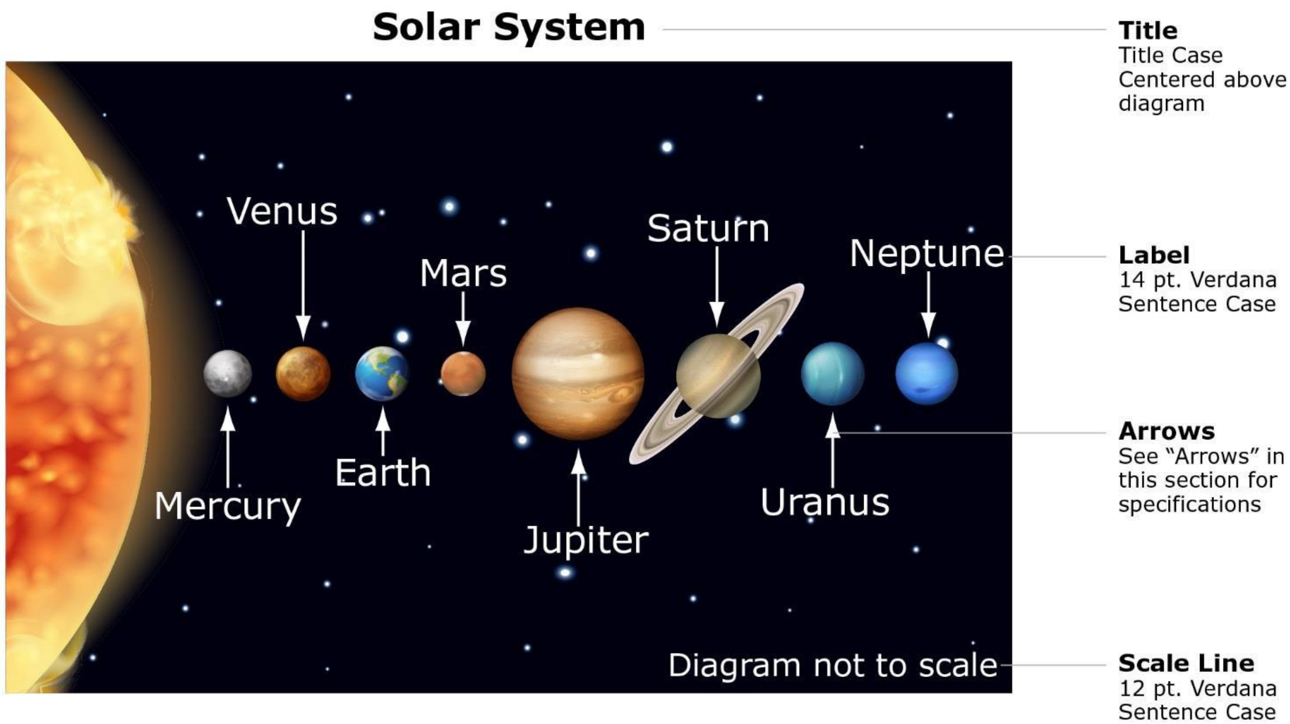
- large enough for students to read text and view content.
- small enough to fit in the viewing area on the computer screen. Students should not have to use horizontal scrolling to see an entire graphic.
- free of excess white space. Condense graphics as much as possible without compromising legibility and font size.

About Image Dimensions
The table below provides the recommended maximum image dimensions (in pixels) you should use for each item layout in order to avoid forcing students to view the item's content on a standard 1024 x 768 monitor. These dimensions should not be regarded as absolute restrictions but as suggested guidelines to consider when inserting images in IAT.

	Stem Max Image Width	Stem Max Image Height	Passage Max Image Width	Passage Max Image Height
Layout 1	600	590		
Layout 2	400	450		
Layouts 3 & 4	400	450		
Layout 5	600	400		
Layout 6	600	400		
Layout 8				
Layout 12				
Layout 13				
Layout 14	600	450		
Layout 15				
Layouts 11 & 17	360	250	590	590
Layout 21	570	250	400	590
Layout 22				
Layout 23	960	250	980	250
Layout 24 & 29	600	250	380	590
Layout 25	670	250	290	590
Layout 26	470	250	500	590
Layout 27 & 28	270	250	700	590
Layout 32 & 33	960	700		
Layout WAI	960	700	960	700

Scaled Graphics

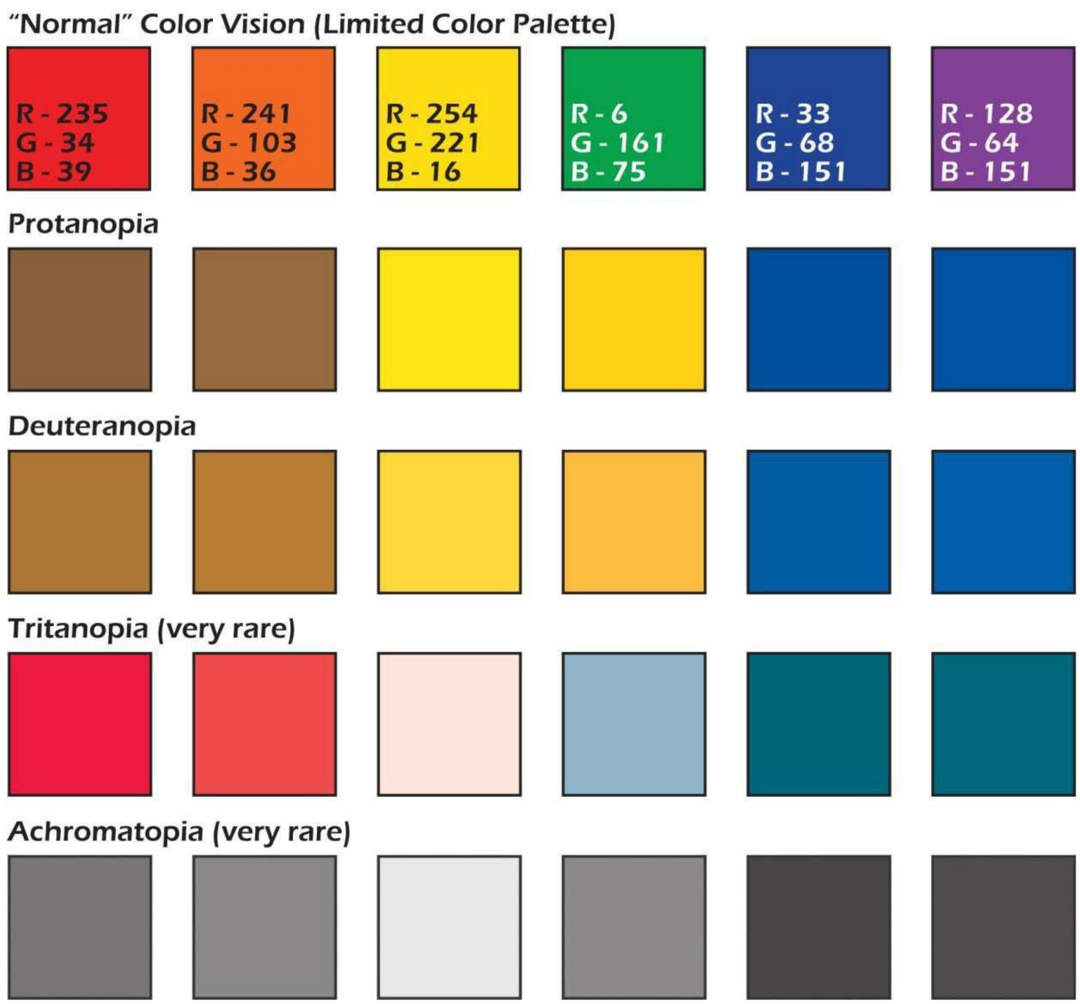
Graphics that are not drawn to the correct scale are labeled with the phrase "Diagram not to scale." Note that "Diagram not to scale" is not punctuated.



C. Graphic Colors

Although color can be used to enhance the appearance of graphics, it should be used sparingly (at the content specialist's discretion). The use of color introduces special considerations for students with color-vision deficiencies. Use the guidelines that follow to help avoid creating graphics that present challenges for these students.

- Do not design graphics that require students to rely solely on color to obtain information.
- Consider combining colors with pattern fills to assist students who might have trouble using color alone to differentiate graphic elements, such as bars on a graph.
- Use a color-vision deficiency simulator, such as Vischeck or Coblis, to check colors in graphics for possible issues.
- Use the limited color palette shown in the top row of the diagram that follows. The other rows in the diagram show how the colors in the limited palette appear to students with certain color-vision deficiencies.



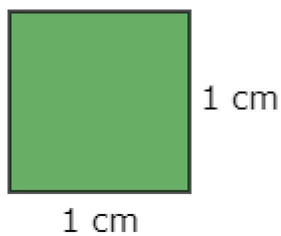
D. Common Graphic Elements

This section provides specifications for elements that often appear in graphics.

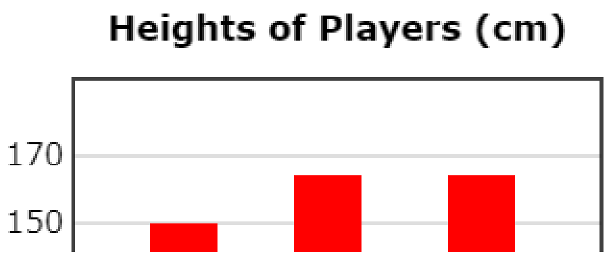
Note: Within individual graphics, the sizes of elements such as points, tick marks, and tallies may be adjusted as needed to emphasize or to de-emphasize certain content in a graphic.

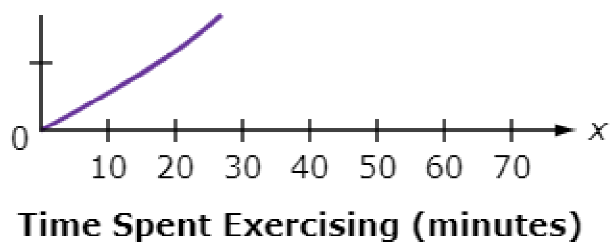
Alignment
Graphics within both the stimulus and the items should be centered on the page.

Units of Measure
In graphics, units are abbreviated or expressed as symbols. (See Part IV for the correct abbreviations of units.) The abbreviated unit or symbol is not preceded by the word *in*.



In graphs and tables, units are enclosed in parentheses after titles or headings.





In tables, do not include the unit in both the column/row heading and in the individual cells in the column/row.

Incorrect:

Shots Made

Game	Shots Made (%)
1	65%
2	70%

Correct:

Shots Made

Game	Shots Made (%)
1	65
2	70

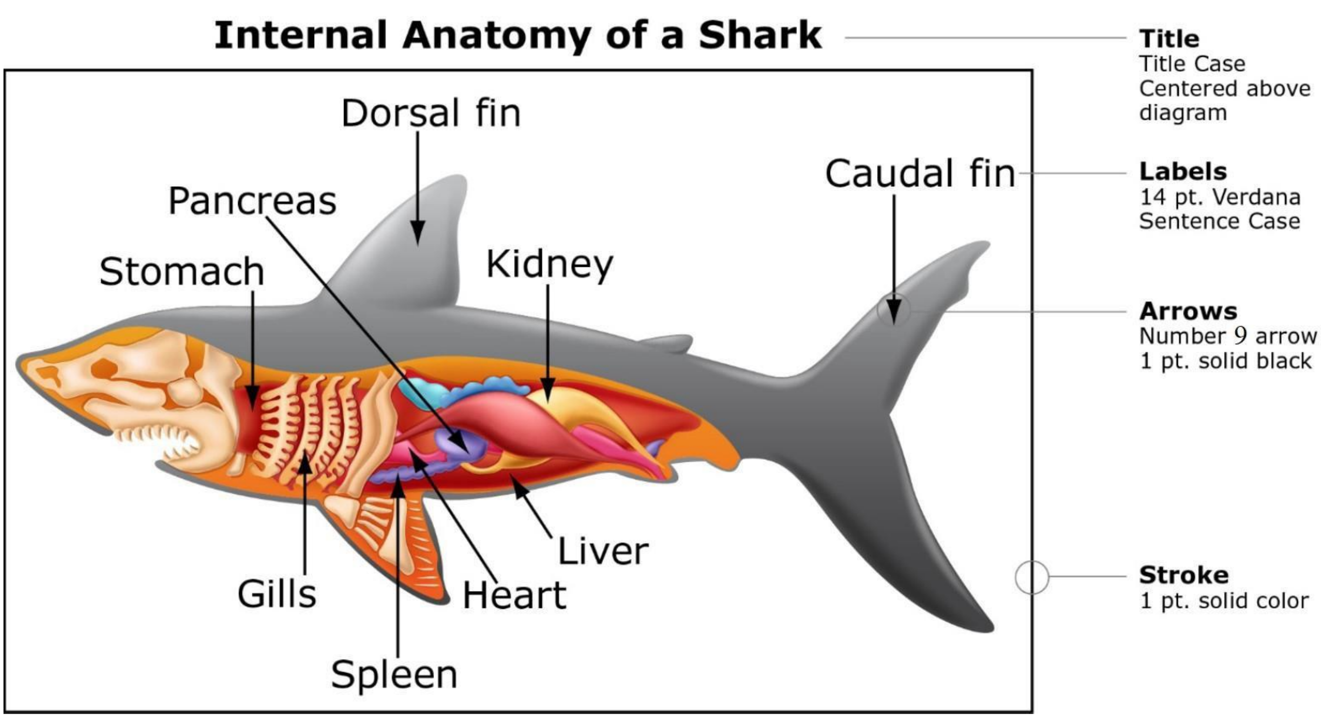
The diagram that follows shows some specifications for diagrams.

- A diagram is art used to describe a scientific system; therefore, the diagram must be scientifically accurate, in realistic perspective, and in scale.
- Tabletops are indicated only by a horizontal line at the back of the table. It is generally not necessary to show the legs or front edge of the table. The tabletop is labeled as "Tabletop" without an arrow.
- Diagrams describing a controlled experiment setup should show the system being investigated with both the manipulated and responding variables.
- All the materials of an experiment do not need to be in the diagram.
- Students, when appropriate, may be included in a diagram but the students must be realistic and grade appropriate.
- Text in diagrams should be phrases, not sentences.
- Any diagram in an item that comes from a diagram in the stimulus should be the same.
- Graphics and diagrams are centered.

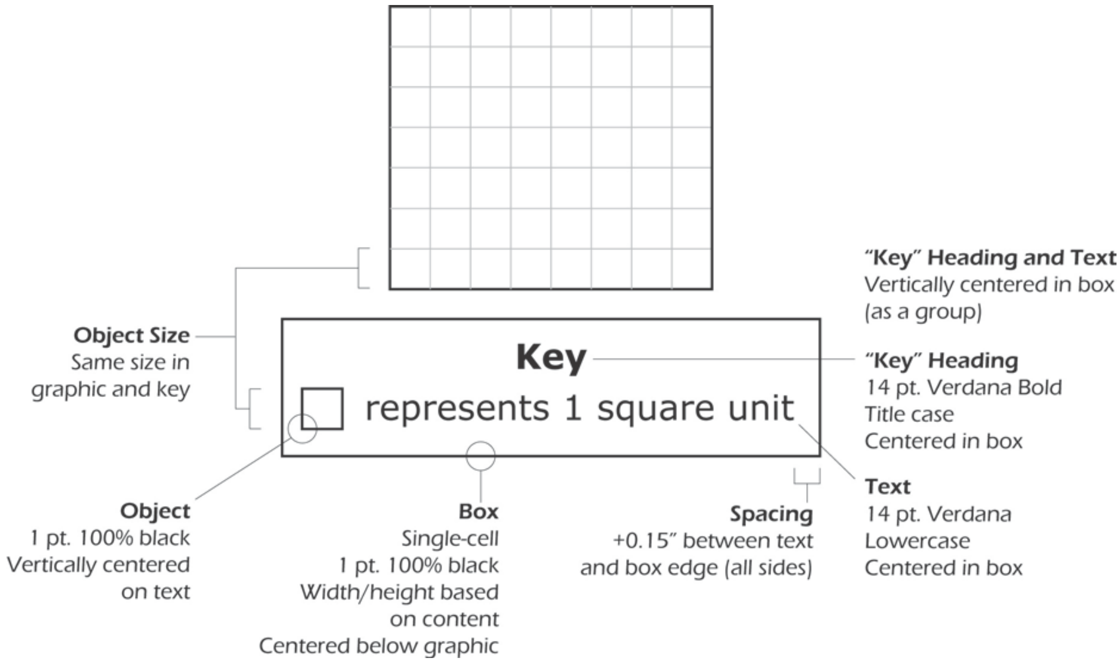
Arrows

The diagrams that follow show specifications for arrows and arrowheads used in graphics to label diagrams.

- Use number 9 arrowhead in Illustrator.
- The arrowhead of the leader from the label to the object being labeled should touch the outer rim of the object being labeled, not inside the object unless what is being labeled is inside an object. For example, a beaker partially filled with a liquid with an arrow labeling the liquid.
- Arrows may be any size or shape as needed in a specific diagram.
- Curvy arrows are no longer used in science assessments. We are still working to pull together some accurate examples.



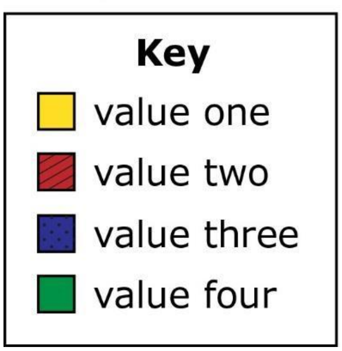
Keys and Scales
Keys and scales are used to provide information that helps students understand graphics. The diagram that follows provides specifications for keys and scales that appear in graphics.



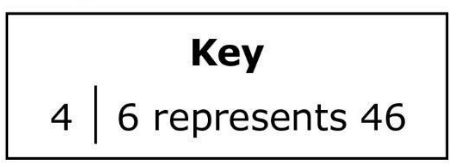
Key vs. Scale
Use a key to provide students with information that helps them identify parts of a graphic or interpret information in a graphic. Use a scale to provide students with ratios and conversions.

In the diagrams that follow, Example 1 tells students the values of graphic elements (e.g., bars in a graph, parts of a shape) filled with colors and patterns; Example 2 tells students how to interpret information in a stem-and-leaf plot; and Example 3 provides students with a ratio for converting centimeters to kilometers.

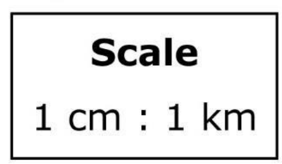
Example 1: Patterns



Example 2: Stem-and-Leaf Plot



Example 3: Scale/Ratio



E. Photographs and Raster Images

Table 9 provides general specifications for scanning photographs and raster images. Certain specifications differ for computer-based testing (CBT) and print-based testing (PBT). Raster images should not be converted to a different form like .svg.

Table 9. CBT and PBT specifications for photographs and raster images

Specifications for Photographs and Raster Images	
Specification	**Requirements**
File format	- CBT: SVG, PNG - PBT: SVG, PNG
Image resolution	- CBT: 72 dpi - PBT: 300 dpi
Color mode	- CBT: RGB - PBT: CMYK
Dimensions	- Based on test content
File Size	- < 1 Mb

Note: Always obtain a high-resolution image that meets the requirements for both CBT and PBT.

Image Improvement
All scanned images should be optimized to represent an ideal tonal range. When scanning images, eliminate any trace of moiré patterns (the line screens that appear when a printed image is scanned). Use the following techniques to help minimize the appearance of moiré patterns:

- Scan images at a 45° angle.
- Set the scanning software to compensate for line screens.
- Apply the Gaussian Blur filter in Adobe Photoshop.

Credit Lines

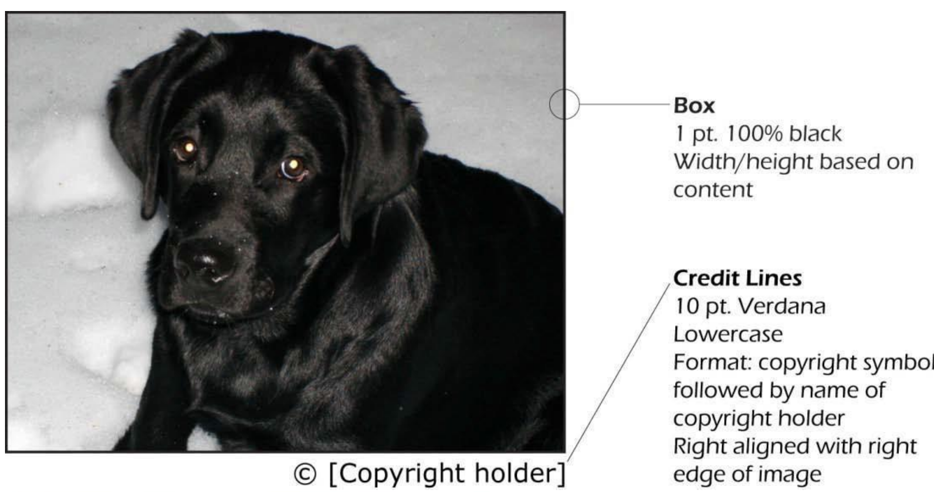
A credit line must be included for all images taken from copyrighted sources and those with Creative Commons licenses, including print publications, the Internet, stock photo agencies or discs, and other commercial and noncommercial sources. For images that do not appear in selections, credit lines should appear as shown in the examples below. (For information about images that appear in selections, see "Acknowledgments" in Part III.) All photographs and raster images should be enclosed in a box, as shown below.

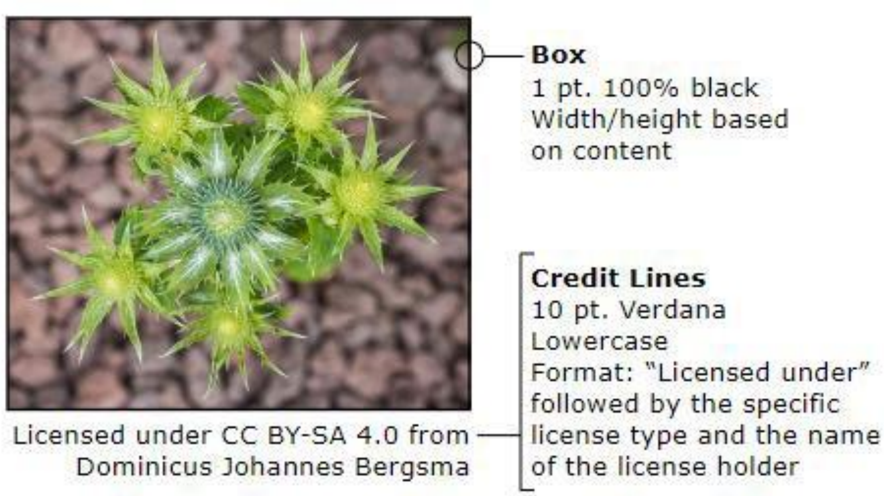
A credit line should be added by the graphics team as a caption that is part of the image. Credit lines should not be done in the IAT Editor.

Copyright is represented by the copyright symbol. Creative Commons licenses are represented by text in the format of "Licensed under CC BY-SA 2.0," indicating the type of Creative Commons license the image has.

For any license before Creative Commons 4.0, the credit line should include the name of the image. In instances where there are multiple images and one is a 4.0 or later and another is 3.0 or earlier, list the image name for both to be consistent within the item.

The Copyright Attribute in ITS should include the original name of the image, the owner of the image/copyright or license holder, a link to the original image, and a link to the copyright or license type.

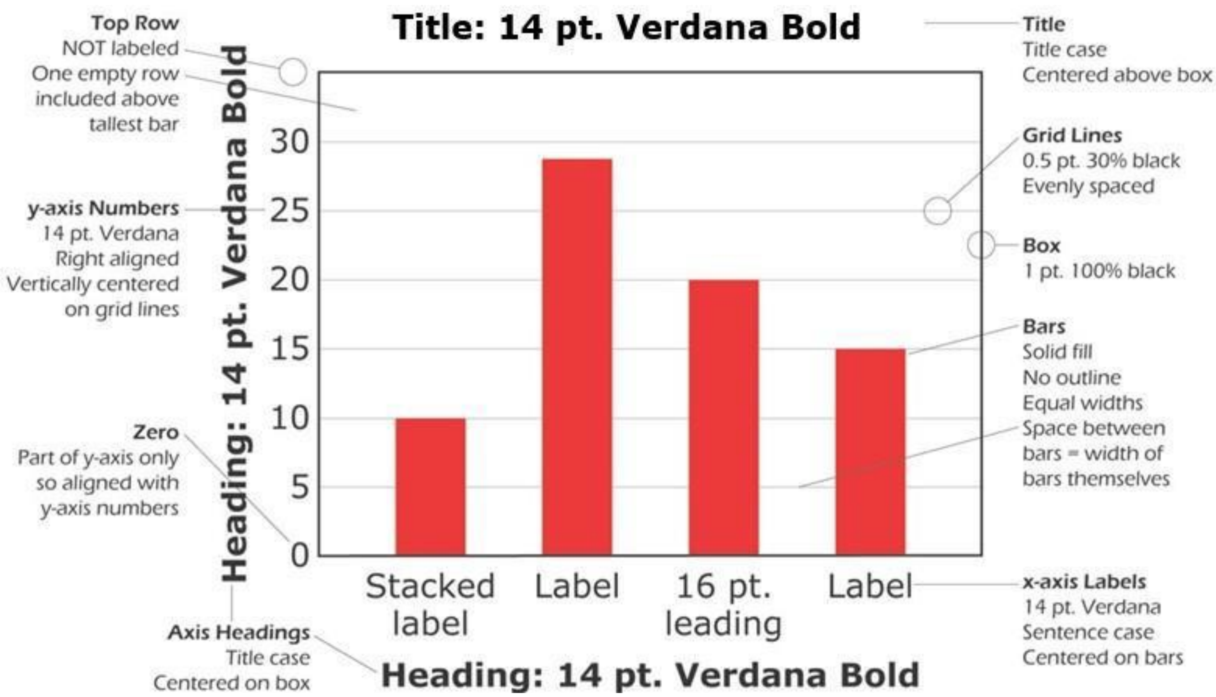


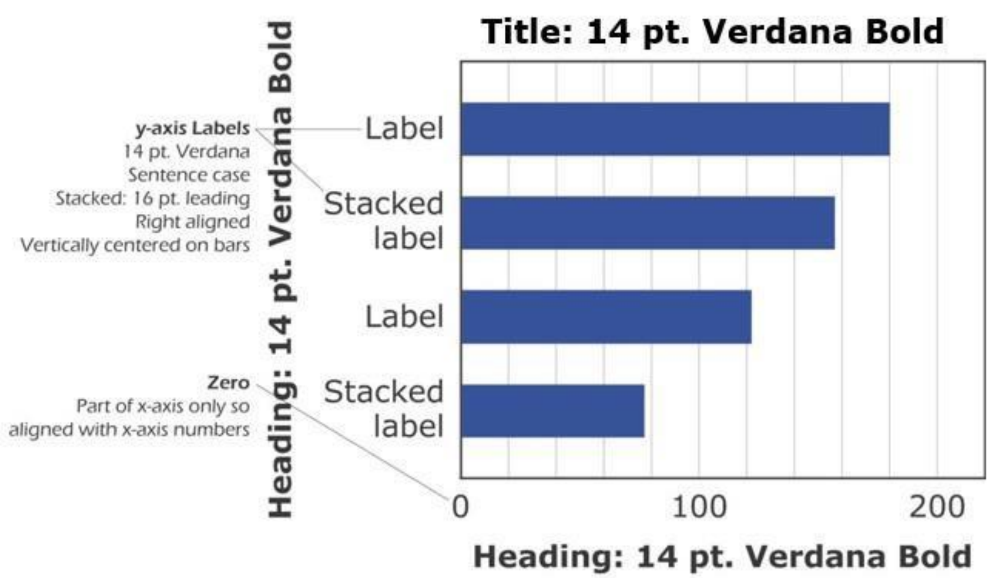


F. Graphs

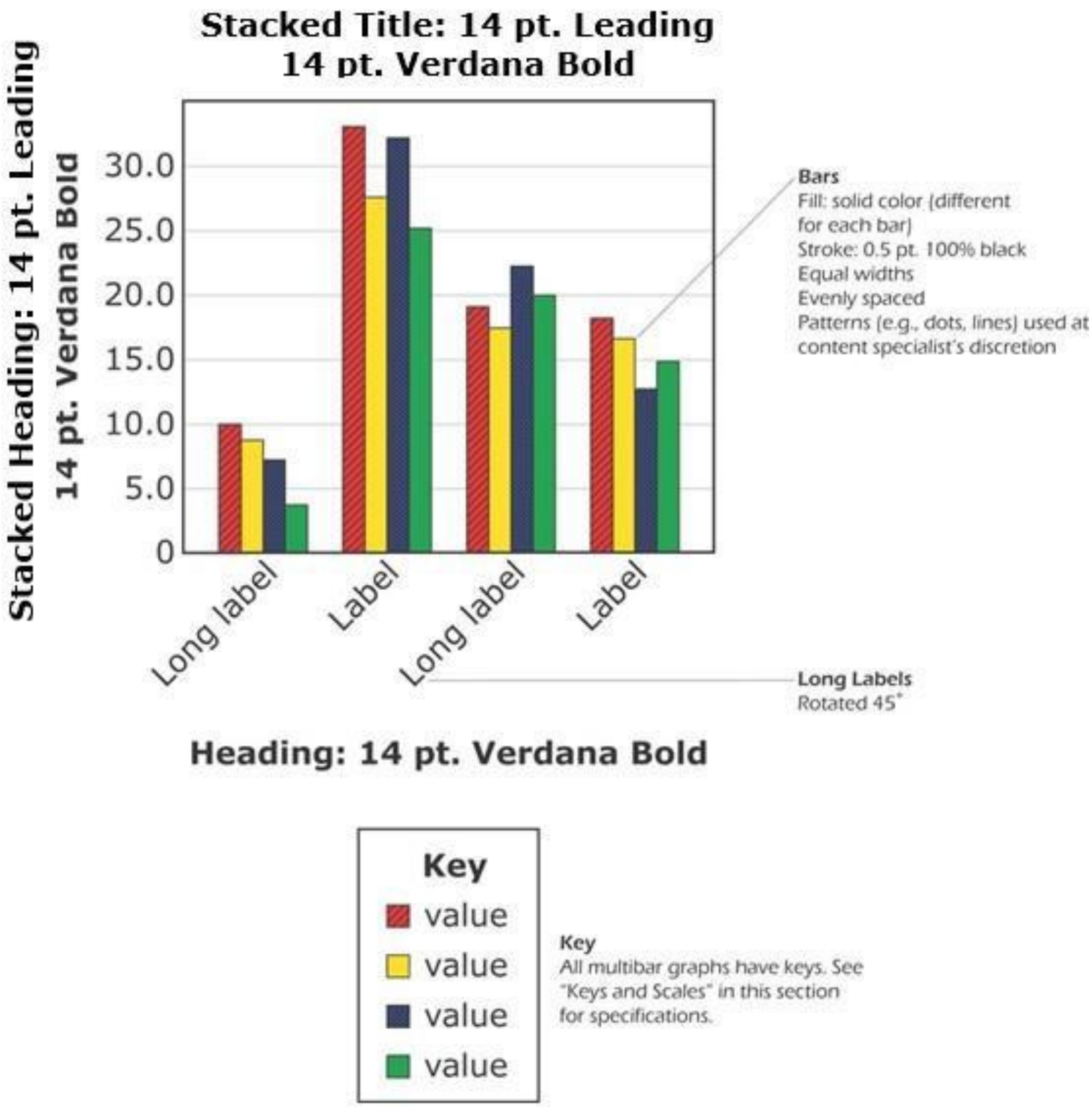
The diagrams in this section provide specifications for bar graphs, histograms, circle graphs, line graphs, and scatter plots.

Bar Graphs

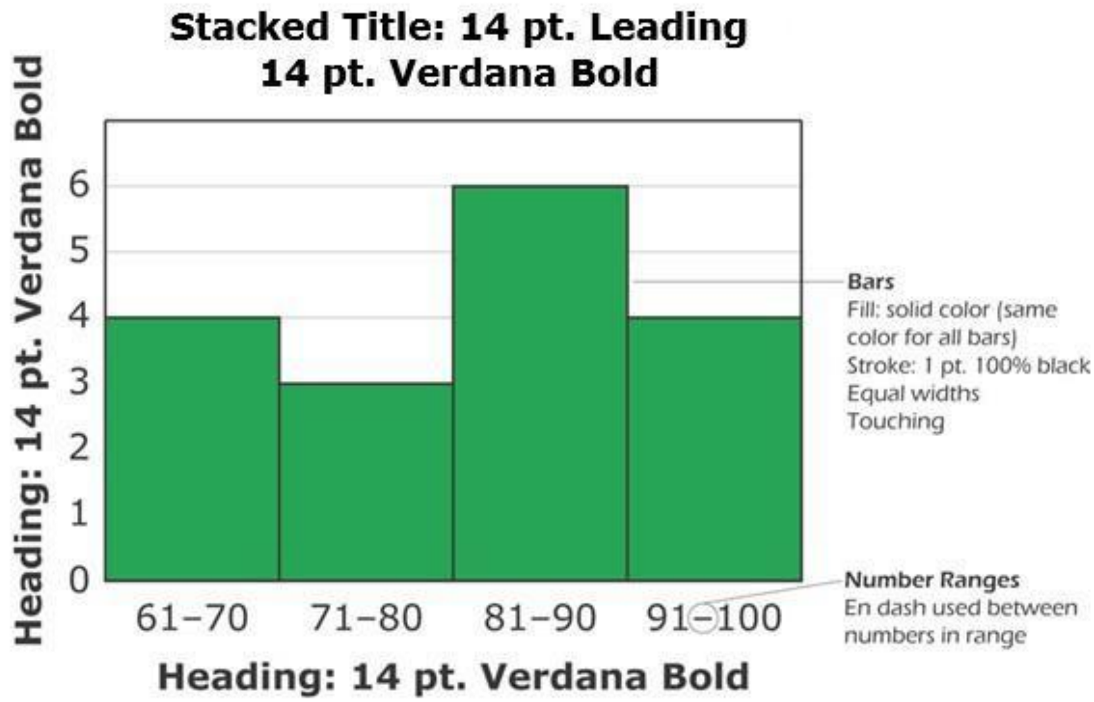




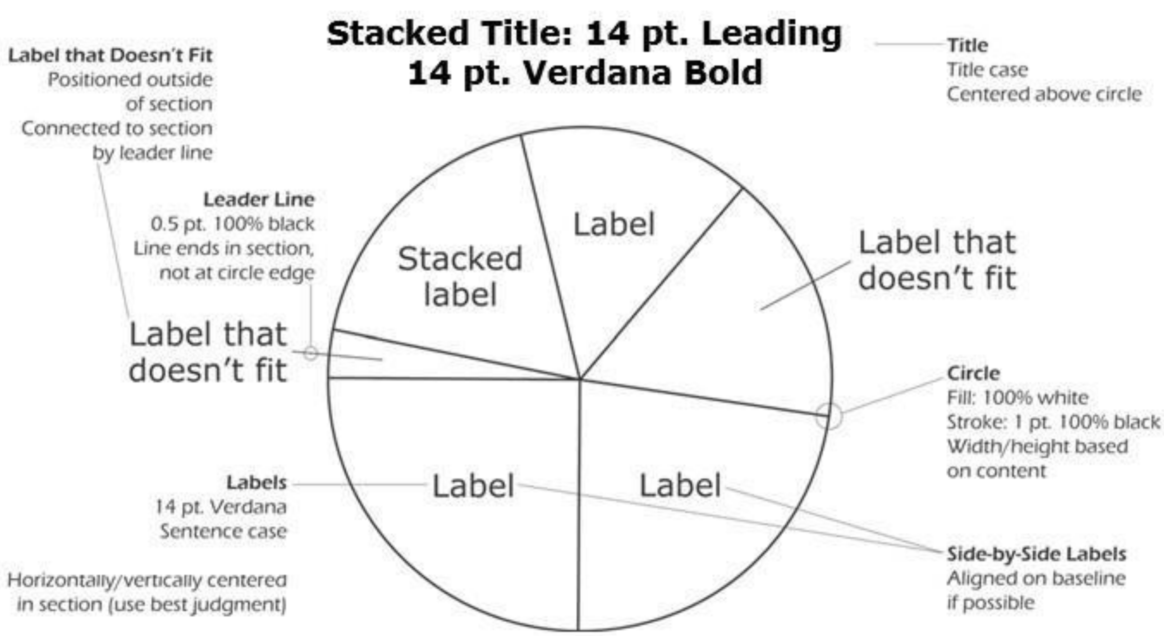
Multibar Graphs



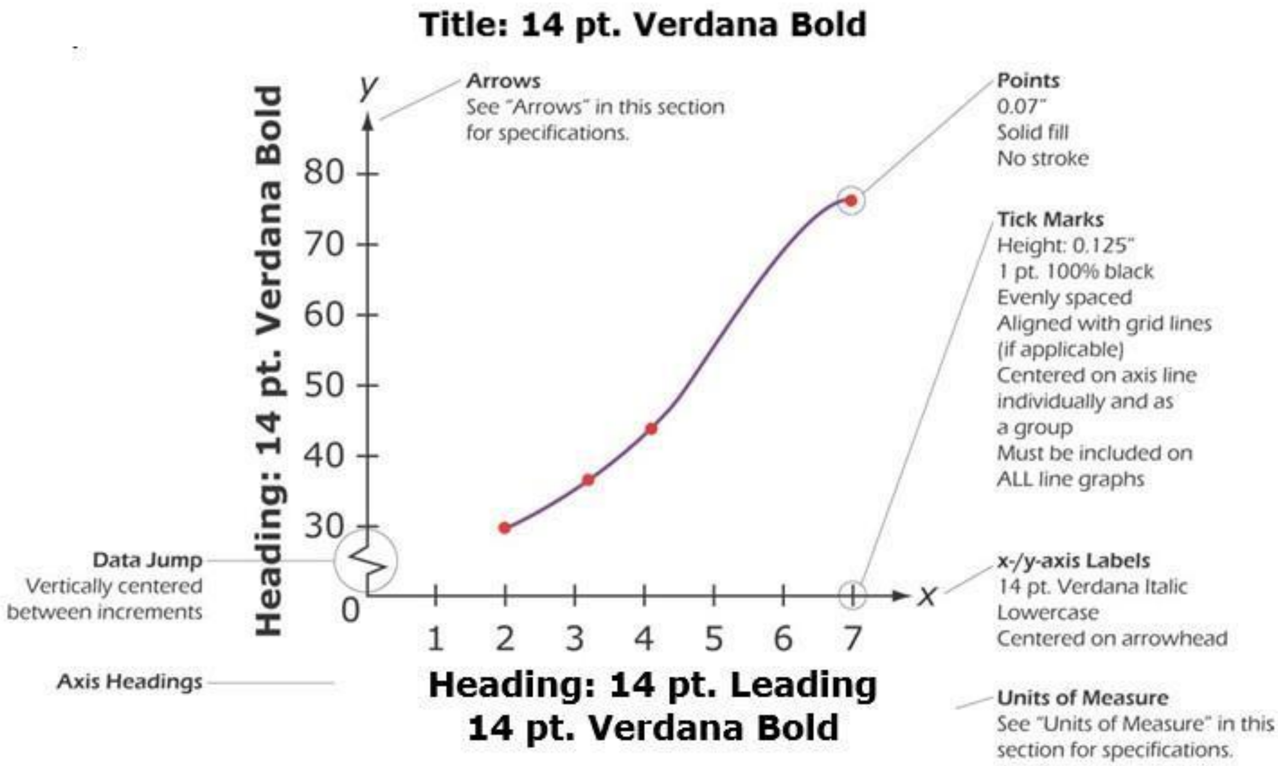
Histograms



Circle Graphs



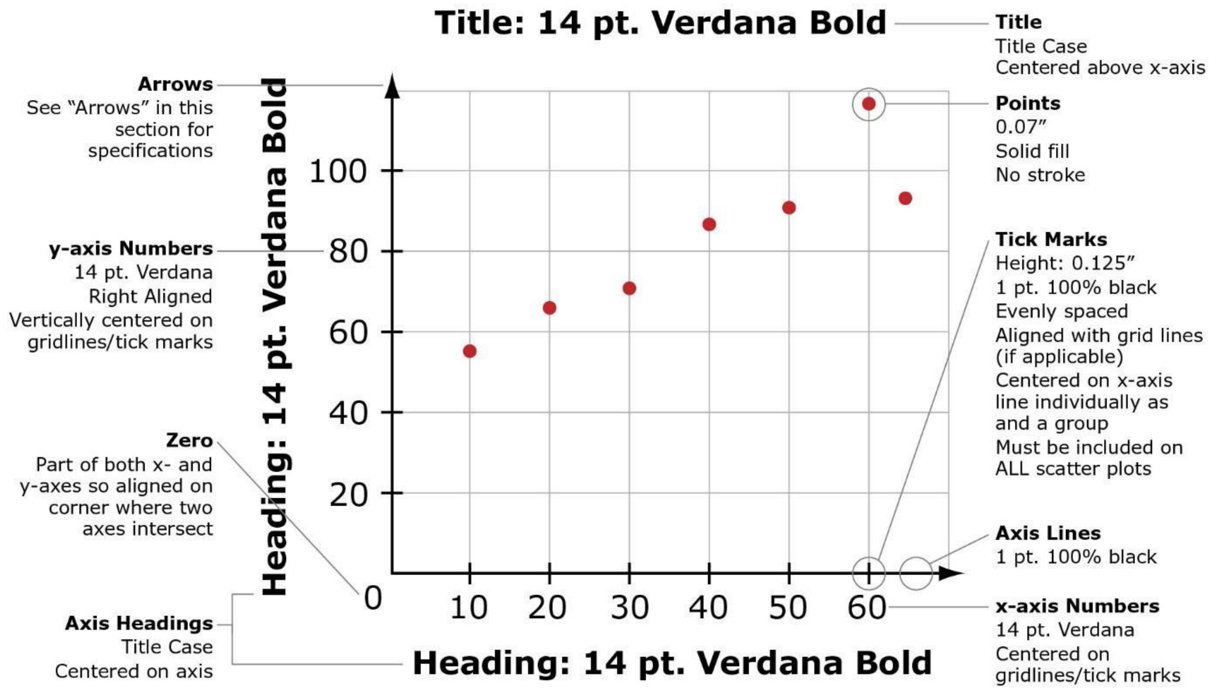
Line Graphs

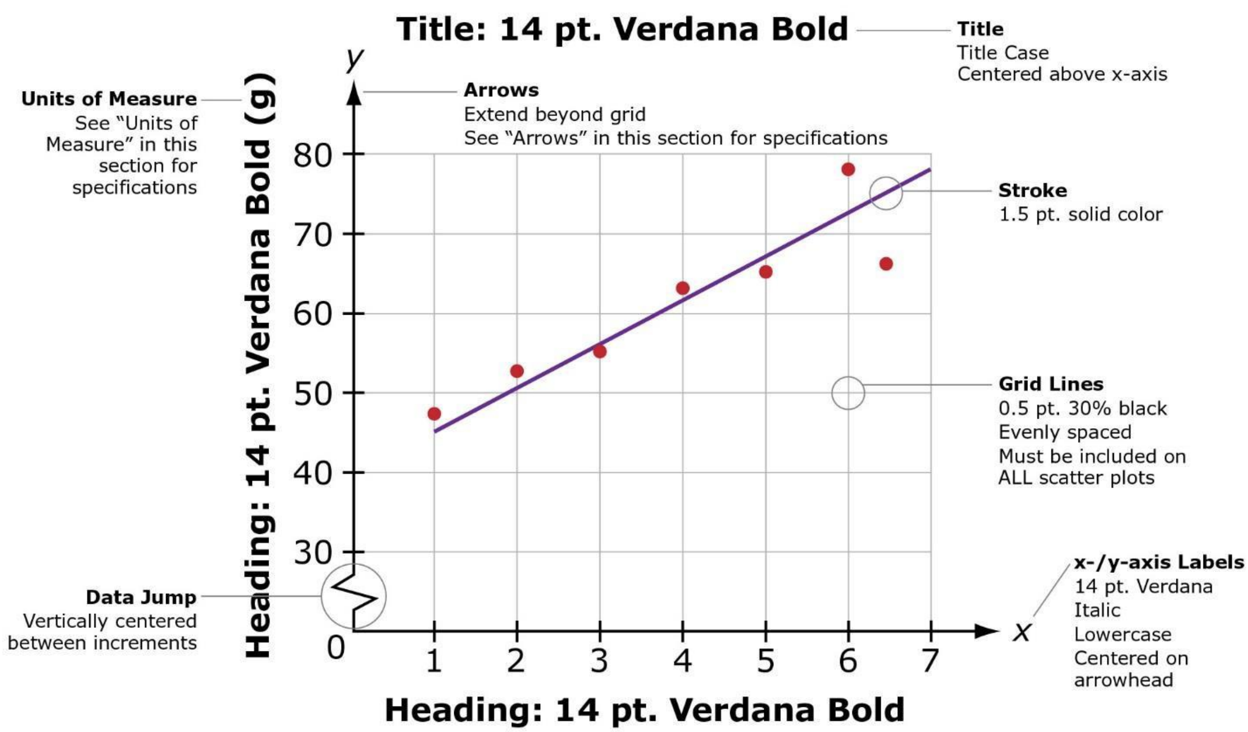


Title: 14 pt. Verdana Bold
Arrows
See "Arrows" in this section for specifications.
Points
0.07"
Solid fill
No stroke
Tick Marks
Height: 0.125"
1 pt. 100% black
Evenly spaced
Aligned with grid lines (if applicable)
Centered on axis line individually and as a group
Must be included on ALL line graphs
Heading: 14 pt. Verdana Bold
y
80
70
60
50
40
30
0
1 2 3 4 5 6 7
x
Data Jump
Vertically centered between increments
x-/y-axis Labels
14 pt. Verdana Italic
Lowercase
Centered on arrowhead
Axis Headings
Heading: 14 pt. Leading
14 pt. Verdana Bold
Units of Measure
See "Units of Measure" in this section for specifications.

Title: 14 pt. Verdana Bold
Title
Title case
Centered above x-axis
y-axis Numbers
14 pt. Verdana
Right aligned
Vertically centered on gridlines/tick marks
120
100
80
60
40
20
0
Stacked Heading:
14 pt. Verdana Bold
Stroke
1.5 pt. solid color
Grid Lines
0.5 pt. 30% black
Evenly spaced
Zero
Part of both x- and y-axes so aligned on corner where two axes intersect
Axis Lines
1 pt. 100% black
10 20 30 40 50 60 70
x-axis Numbers
14 pt. Verdana
Centered on gridlines/tick marks
Axis Headings
Title case
Centered on axis
Heading: 14 pt. Verdana Bold

Scatter Plots





G. Tables and Charts

The diagrams in this section provide specifications for tables and charts, including pictographs and tally charts.

Text Alignment in Tables

Left align text within a cell when there is more than one word.

Table 1. Title

Text entry here	
Longer text entry here	
Each sentence cased	

Center text within a cell when there is:

- a single word or number
- a combination of words and numbers
- a combination of fractions and mixed numbers

Fractions and Mixed Numbers
Centered

Fractions	Mixed Numbers
$\frac{1}{3}$	7
$2\frac{3}{5}$	$\frac{1}{4}$
$\frac{3}{16}$	300.00

In columns with numbers only, center the widest number and align other numbers on the ones place or decimal point.

Table 2. Title

Longest number centered	100	52.8
Rest aligned on ones, decimal, or symbol	2	325.25
	22	1.5

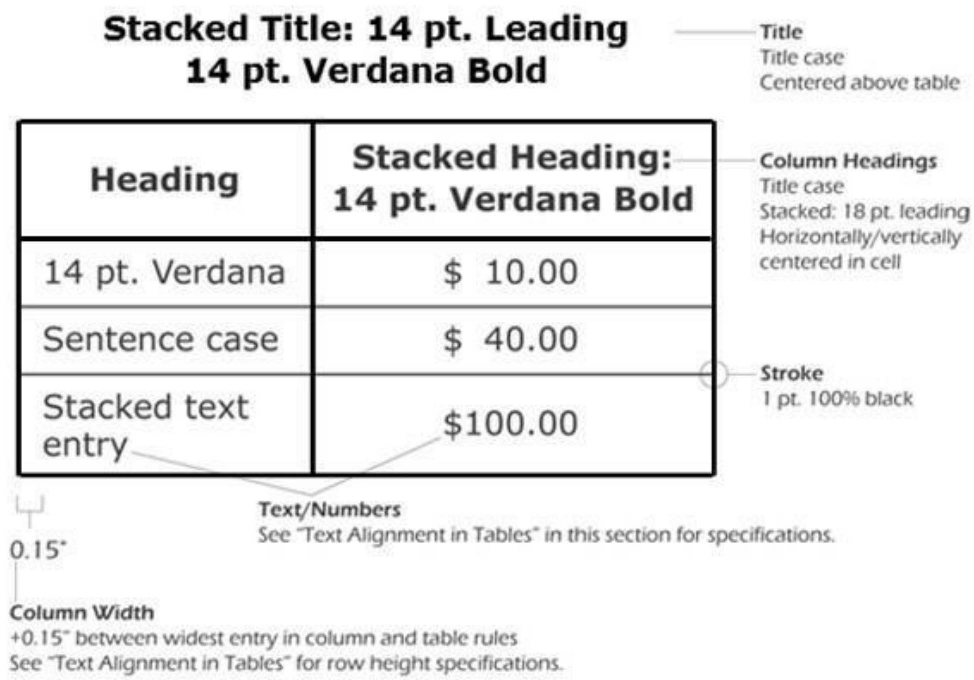
In columns with currency only, center the widest value, then:

- Align the dollar signs.
- Align dollar amounts on the decimal point.

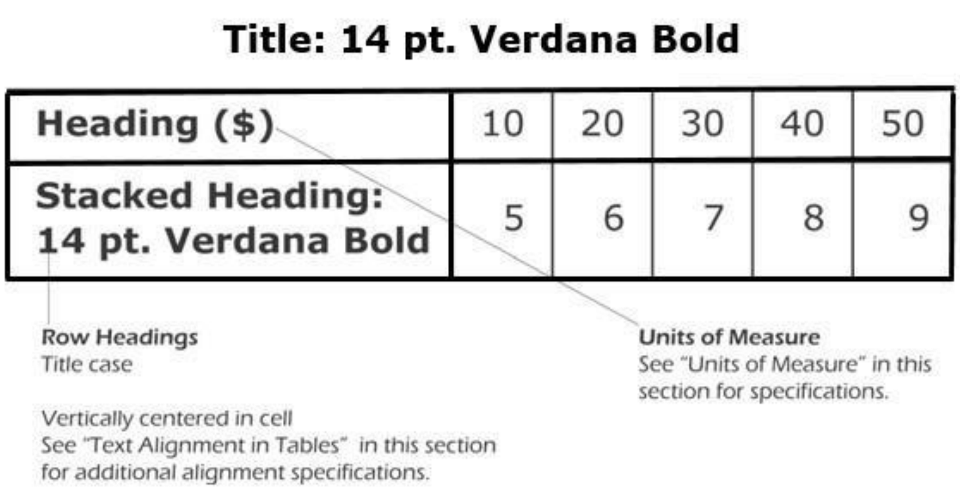
Table 1. Title

Dollar signs left-aligned	$12,540.00
Dollar amounts aligned on the decimal point	$ 25.00

Tables

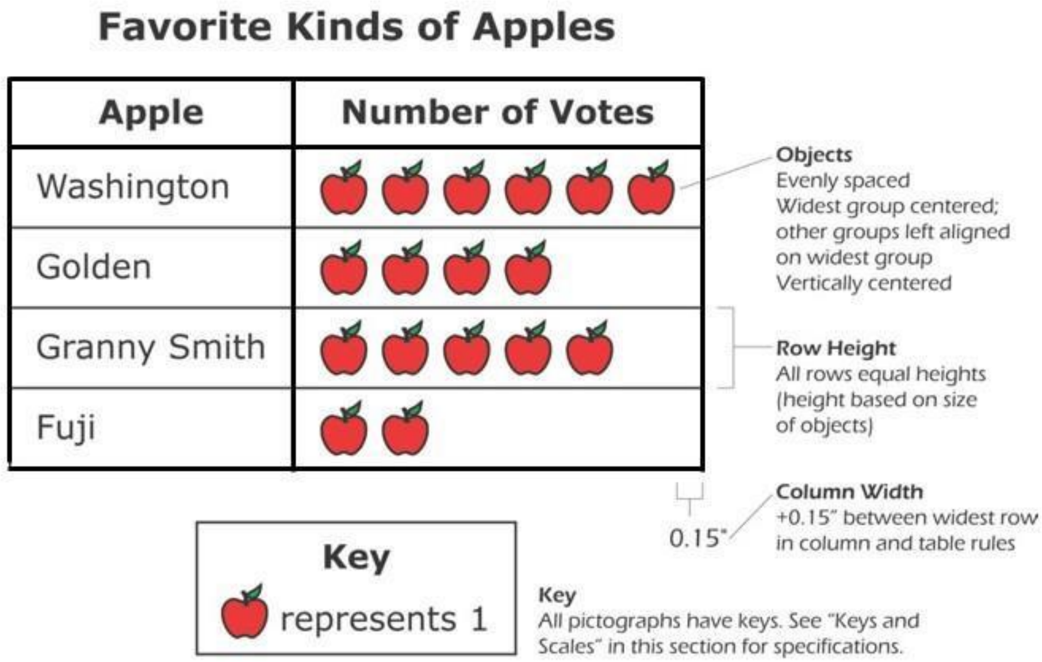


Note: If a cell entry begins with a numeral followed by text, lowercase the first word of the text, just as if the number were spelled out, e.g., use "10 grams," as you would if it were "Ten grams."



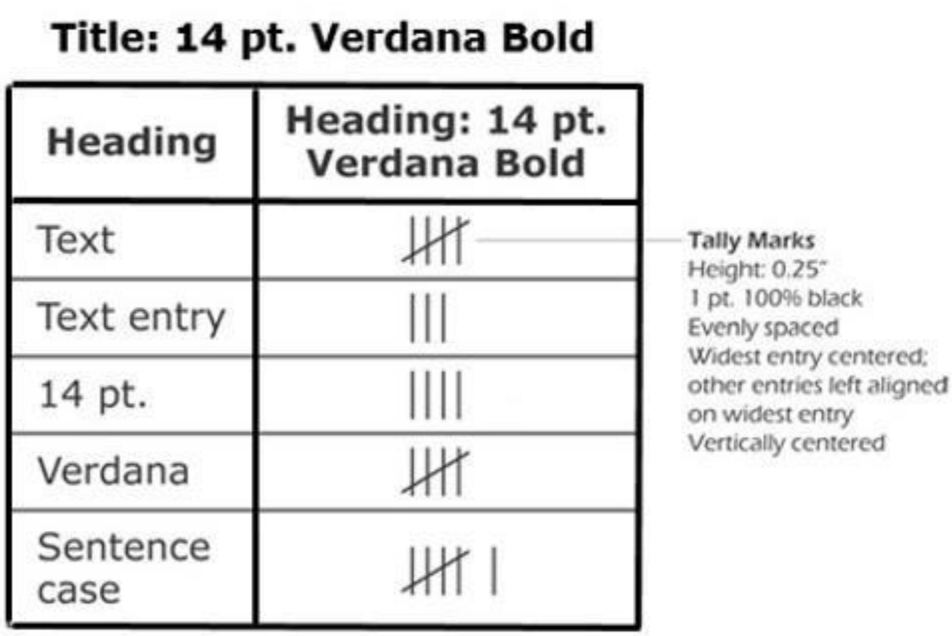
NOTE: ALL TABLES, EXCEPT FOR MATCHING ITEM TABLES, REQUIRE TITLES. TABLES ARE ALL CENTERED.

Pictographs



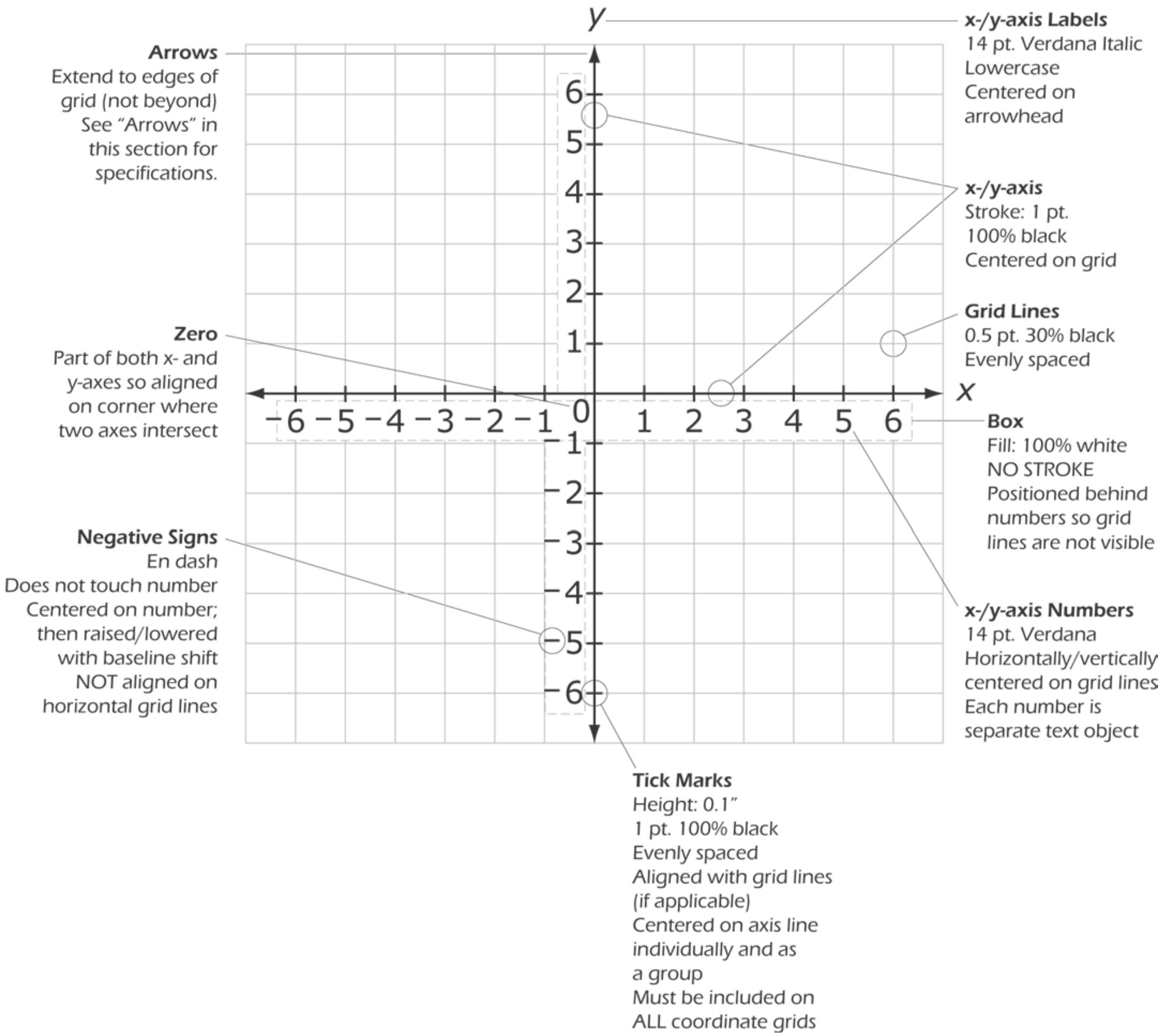
Note: People and animals should NOT be represented as half symbols in pictographs.

Tally Charts

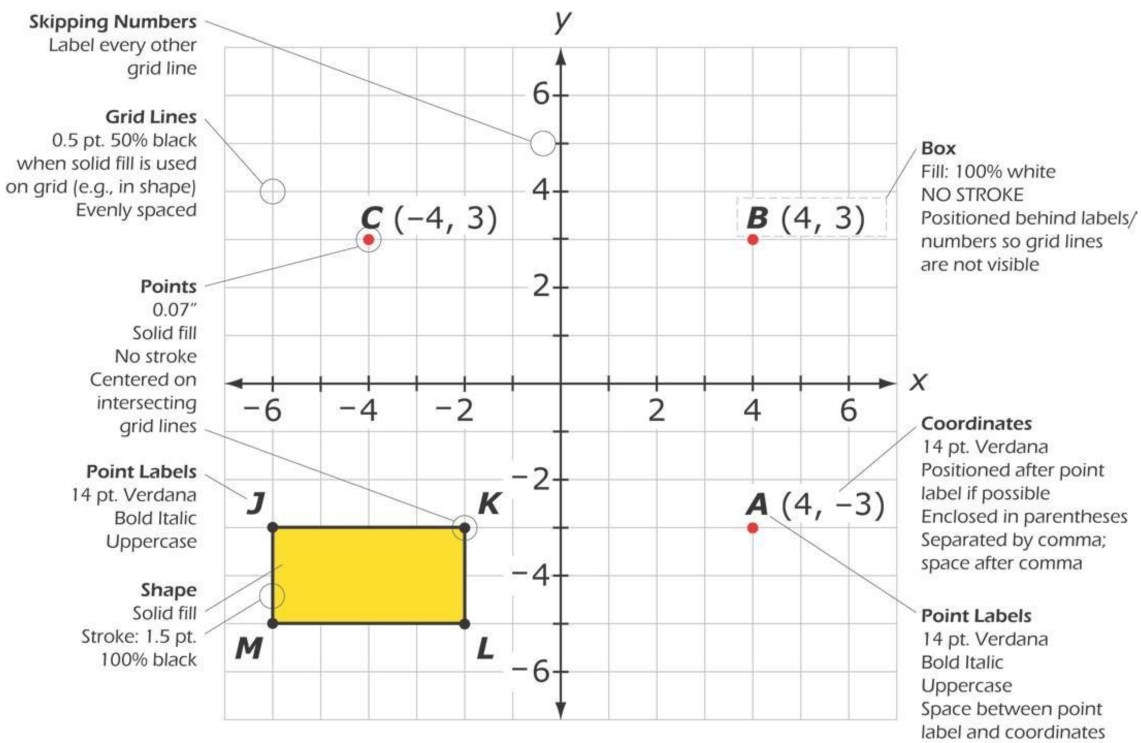


H. Coordinate Grids

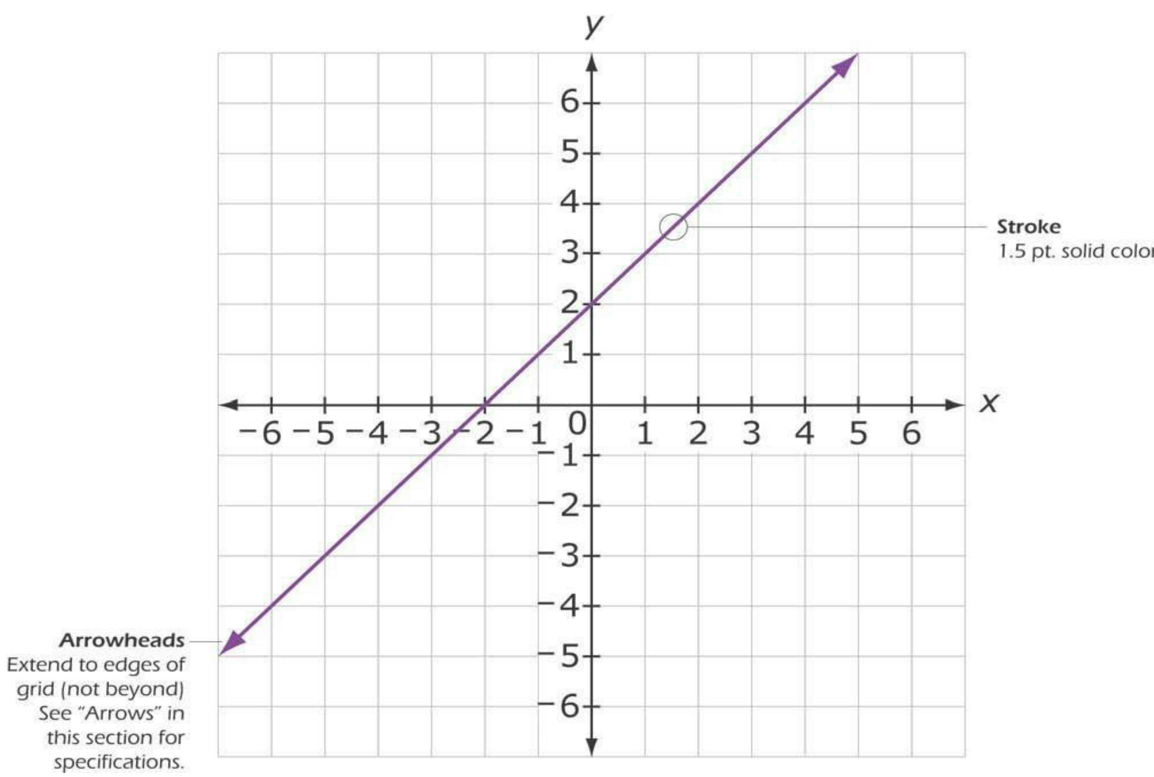
The diagrams in this section provide specifications for coordinate grids, including coordinate grids with plotted points, lines, and shapes.



Plotted Points and Shapes



Plotted Lines

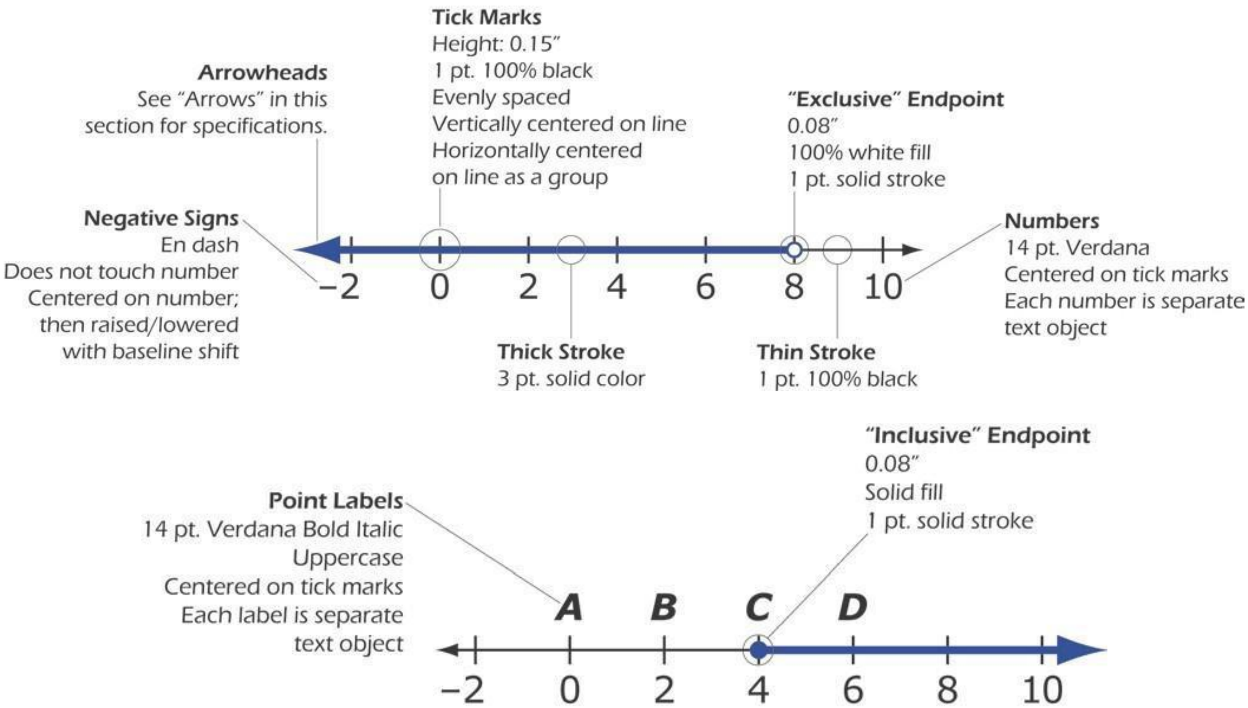


Note: Use graphing software to ensure that lines are accurately plotted on coordinate grids.

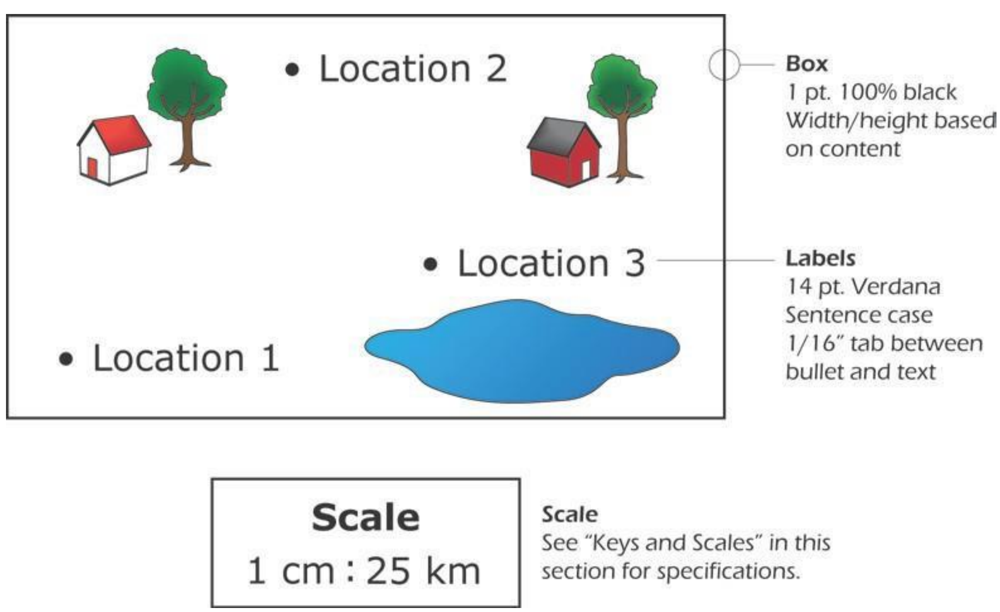
I. Other Types of Graphics

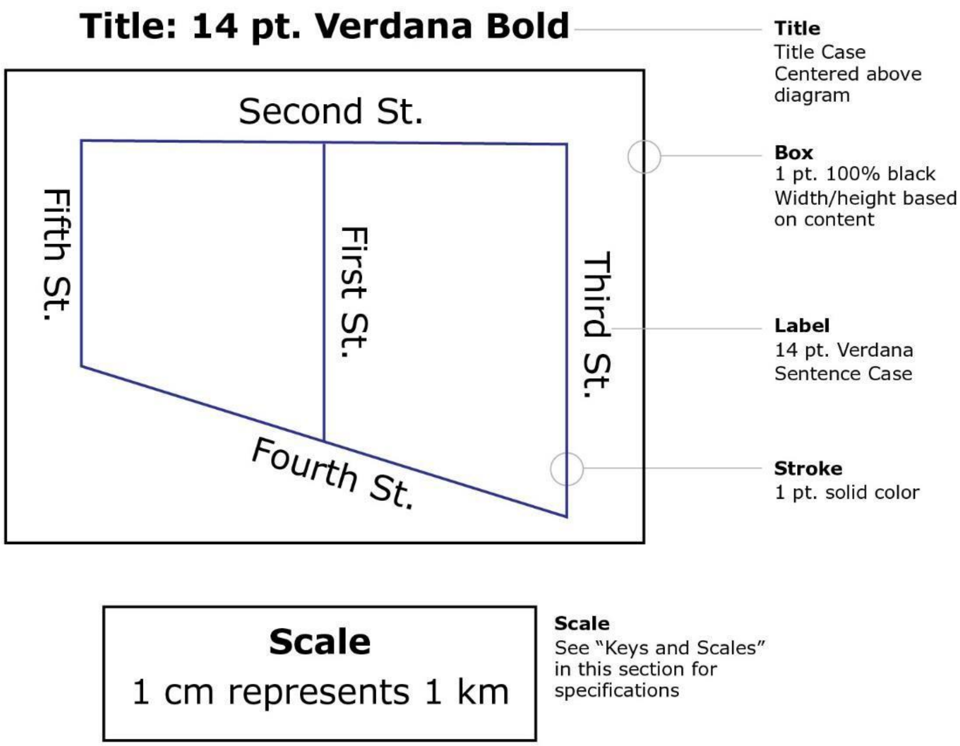
The diagrams in this section show specifications for miscellaneous graphics, including number lines, box-and-whisker plots, line plots, stem-and-leaf plots, spinners, maps, and graphic organizers.

Number Lines



Maps




Title: 14 pt. Verdana Bold
Second St.
Fifth St.
First St.
Third St.
Fourth St.
Title
Title Case
Centered above diagram
Box
1 pt. 100% black
Width/height based on content
Label
14 pt. Verdana
Sentence Case
Stroke
1 pt. solid color
Scale
1 cm represents 1 km
Scale
See "Keys and Scales" in this section for specifications

J. Animations

- Graphics for animations and simulations follow three-dimensional science graphics style guidelines.
    - Animation art style should not be different from those produced for APG static images.
- Dimensions for animations and simulations:
    - In a stimulus: animation can be no wider than 350 pixels to eliminate horizontal scrolling.
    - In a simulation: dimensions are predetermined by the item layout.
- Animations, graphics, and simulations must have file sizes that keep the entire item or cluster's file size under 500 kb.
- Animations and simulations should be short and succinct as to reduce the amount of test time spent running them.
- For animations (not simulations), include the directive "Click on the small gray arrow to start the animation …" OR "Click on the small gray arrow to watch [Earth revolve around the sun, the rabbit come out of its hole, etc.]."

Appendix A: Word List

Word List

This appendix provides guidelines for the capitalization, hyphenation, and spelling of terms. See "Compound Terms" in Part I for additional guidelines.

A
AD (uppercase; precedes date)

a.m. (lowercase)

B
BC/BCE (uppercase; follows date)

big bang theory (lowercase)

C
criterion (singular), criteria (plural)

cutout (n, adj), cut out (v)

D
data are (plural)

E
Earth, in reference to the planet (not used with the article the)

earth, in reference to earth material (used with the article the)

eastern hemisphere (lowercase)

equator (lowercase)

F
fall (season)

Force is greater, not larger: According to Sir Isaac Newton's Law of Universal Gravitation, all objects that have mass are attracted to each other. Mass is the measure of an object's matter (what it's made up of). The greater an object's mass, the greater its gravitational force.

flowchart (per Webster's)

fresh water (n), freshwater (adj)

full-size (adj)

H
the Hawaiian Islands, but the island of Hawaii

high-pressure (adj), high pressure (n)

K
kinetic energy (KE)

L
landfall

Law – uppercase in terms like “Newton’s First Law” or “Coloumb’s Law”

life cycle

life-span

lightbulb

M
moon

N
northern hemisphere (lowercase)

P
potential energy (PE)

R
rain forest (n)

round-trip (n, adj)

S
salt water (n), saltwater (adj)

setup (n), set up (v)

solar system

southern hemisphere (lowercase)

spring (season)

summer

sun

T
tide pool

tidewater

timeline

U
underwater (adj, adv)

W
western hemisphere (lowercase)

Appendix B: Resources, Research, and Bibliography

Recommended Resources

The style conventions and specifications in this document are based largely on information from the sources listed in this section. Refer to these sources for additional information.

Chicago Manual of Style and Words into Type
First published in 1906, the Chicago Manual of Style (CMOS) is one of the oldest and most comprehensive editorial style guides available. The sixteenth edition, published in 2010, was updated in consultation with a broad range of scholars and professionals in the fields of academics and publishing.

Different style guides have different uses. For example, the Associated Press Stylebook is used primarily by journalists, the Publication Manual of the American Psychological Association is used for social science publications, and the MLA Handbook for Writers of Research Papers is most commonly used by writers of research papers in the humanities. CMOS is widely used in educational, scholarly, and trade publishing, as is Words into Type, another well-respected general-purpose style manual. Both CMOS and Words into Type provide broader coverage of mechanical issues, grammar, and usage than the more specialized style manuals listed above.

Merriam-Webster's Collegiate Dictionary
Because hundreds of new words are added to the English language each year, and preferences regarding issues such as spelling and hyphenation change over time, it's important that publishing professionals use a recent edition of a good dictionary. Merriam-Webster's Collegiate Dictionary is not only one of the most popular dictionaries on the market (as the best-selling dictionary in the United States) but is also recommended by CMOS and used by many educational and academic publishers.

Garner's Modern American Usage
Bryan A. Garner wrote the grammar-and-usage chapter of CMOS, and his usage book is an appropriate companion to CMOS, significantly expanding on the usage guidelines provided in chapter 5 of CMOS. Garner's Modern American Usage is considered one of the best books available on contemporary usage. Garner takes a prescriptive approach, which means that his usage guidelines are based on established norms and rules for the way language should be used.

The Copyeditor's Handbook: A Guide for Book Publishing and Corporate Communications
Amy Einsohn, the author of The Copyeditor's Handbook, has twenty years of experience as a professional editor and teacher of copyediting classes and is highly regarded in the copyediting community. Her book, which addresses the ABCs of copyediting, editorial style, and language editing, is intended to be used as a tool for self-instruction or a textbook for copyediting classes. The Copyeditor's Handbook has been recommended by current and former editors at CMOS, including Carol Fisher Saller and Margaret Mahan.

Support for Selected Style Preferences

Emphasis Terms

The treatment of various elements of style, including emphasis terms, varies according to the preferences of the user, but it is recommended that emphasis terms in test items be set in boldface. The publication Considerations for the Development and Review of Universally Designed Assessments[1] cites the following arguments for and against the various treatment options for emphasis text:

- Standard typeface, uppercase and lowercase, is more readable than italic, slanted, small caps, or all caps (Tinker, 1963).
- Text printed completely in capital letters is less legible than text printed completely in lowercase, or normal mixed-case text (Carter, Dey, and Meggs, 1985).
- Italic is far less legible and is read considerably more slowly than regular lowercase (Worden, 1991).
- Boldface is more visible than lowercase if a change from the norm is needed (Hartley, 1985).

Ten style guides from Smarter Balanced states and the consortium were evaluated. Seven of the ten recommended the use of boldface for emphasis terms, whereas one recommended underscoring, another recommended small caps, and the last did not specify. Designers of online content agree that "[o]n the Web, the most common and effective method [for emphasizing text] is the use of a bold face from the current font family."[2]

In addition, the application of a particular treatment to a style element could prove confusing if repeated to represent another style element: In the Smarter Balanced Style Guide, underscoring is used to designate vocabulary terms. That, combined with the fact that the use of two treatment styles for a single element can appear overpowering, reinforces the recommendation that boldface alone be used for emphasis terms. This recommendation is also supported by the Test Accessibility and Modification Inventory (TAMI).[3]

Exponents and Superscript

These elements should be smaller than running text and should not be separated from the preceding text by a space. It is recommended that exponents and superscript characters be smaller than running text (scaled to 70% of base print size). Increasing the size of running text around such characters is not feasible, and increasing the size of the characters themselves introduces a risk that students will misinterpret the characters as running text and not as exponents or superscripts.

---

1 Thompson, S. J., Johnstone, C. J., Anderson, M. E., and Miller, N. A. (2005, November). Considerations for the Development and Review of Universally Designed Assessments (Center on Educational Outcomes Tech. Rep. 42). Retrieved February 6, 2012, from www.cehd.umn.edu/nceo/OnlinePubs/Technical42.htm

2 Hume, A. (2005, December). The Anatomy of Web Fonts. Retrieved February 13, 2012, from //www.sitepoint.com/anatomy<web<fonts

3 Beddow, P. A. (2009). Test Accessibility and Modification Inventory: Quantifying and Improving the Accessibility of Tests and Test Items. Presented at the CCSSO 2009 National Conference on Student Assessment. Retrieved February 6, 2012, from http://peabody.vanderbilt.edu/Documents/pdf/PRO/TAMI_CCSSO_Beddow.pdf

The general accessibility guidelines include discussion of magnification tools. In addition, the accessibility spoken/audio business rules provide guidelines on how mathematical notation is to be presented in spoken form. The magnification tools and spoken support should allow students with visual impairment to access all the information available on screen.

In regard to ordinals, it is recommended to use 1st, 2nd, 3rd, etc., rather than 1^{st}, 2^{nd}, 3^{rd}. In this way, the exponent size consideration is avoided, and ordinals are as easily read as other running text on a page.

Typeface: Verdana

It is recommended to use Verdana for onscreen testing materials, for its readability as compared to Times New Roman and Arial. Although other fonts are available that are specially designed to further enhance readability, these custom fonts may not be as widely available on student computers and may require the test delivery system to supply the font as part of system installation.

Serif fonts (e.g., Times New Roman), which are popular in print, can appear pixilated and blurred onscreen. In contrast, "the straight, low contrast, open strokes of a sans-serif font, such as Verdana, will always leave a good impression on-screen."[4] Verdana, which was designed for the screen, offers a generous amount of white space both between and within (glyphs) the characters. Currently, it is the most commonly used font on the Web, owing to its marked legibility on screen.

Preparation of Materials for Persons Who Have Color Vision Deficiencies

Color is one of the most important aspects of visual communication and can be employed to generate interest or to communicate ideas or feelings. Yet colors for an audience with members who have color discrimination problems should be selected carefully to avoid conveyance of unintended meaning. This is especially true in educational and testing materials. Many of these materials rely on good color perception for the interpretation of graphs, charts and illustrations. Yet even the most carefully thought-out graphic may lead the user to an incorrect answer because of poor color selection.

- **Select colors carefully.** Besides black and white, most color blind individuals can only see two colors, blue and caramel (golden brown). Red, yellow, orange, and green take on shades of caramel; purple takes on shades of blue when viewed by a person with colorblindness.
- **Less is more.** Too many colors used thoughtlessly can confuse and negate the message of a graphic. Settle on four or fewer colors and stick with them. Black and white are counted as colors when designing graphics, even though they are not usually considered colors when talking about vision.
- **Use contrasting colors.** Contrast is an important influence on the legibility of graphics, especially for persons with color discrimination problems. Substantial contrast, i.e., the use of dark values with light values, between the color of the foreground and the background should be employed. High contrast makes materials easier to read by both persons with colorblindness and those with typical vision. Light letters on a dark background or dark letters on a light background are most legible, but remember the actual colors of those combinations are important.

Contrasting Colors Appropriate for Persons with Color Perception Difficulties (in order of best contrast value)

- Use black and white.
- Use dark blue and white.
- Use black and bright yellow.
- Use dark blue and bright yellow.
- Use dark brown and white.
- Use pale blue and black.
- Use yellow and purple.

Notice that yellow is recommended as a common color for graphics to be used by persons with poor color discrimination. This is because yellow maintains luminance longer than any other color. Even though it is perceived as a light caramel color by persons with color blindness, it holds its brightness longer than any other hue, and therefore maintains its contrast when paired with a dark color.

Color Combinations to Be Avoided

- Avoid gray with any color, even another value of gray.
- Avoid red with any color except white or blue.
- Avoid green with any color except white.
- Avoid brown with any color except white or blue.
- Avoid purple with any color except yellow or white.
- Avoid orange with any color except blue or white.
- Avoid two values of the same color, such as light blue and dark blue.
- Avoid a neutral color with any other neutral color.

The importance of proper attention to color selection cannot be overlooked when developing tests for individuals or groups that have color vision or color perception deficiencies.

Source:

Allman, C. B. (2009). *Making tests accessible for students with visual impairments: A guide for test publishers, test developers, and state assessment personnel.* (4th ed.). American Printing House for the Blind.

Bibliography

Apple Publications Style Guide. (2009, December). Retrieved from https://developer.apple.com/library/safari/#documentation/UserExperience/Conceptual/APStyleGuide/ APSG_2009.pdf.

Barr, C., and Yahoo! Senior Editors. (2010). *Yahoo! Style Guide: The Ultimate Sourcebook for Writing, Editing, and Creating Content for the Digital World*. New York: St. Martin's.

Chicago Manual of Style (16th ed.). (2010). Chicago: University of Chicago Press.

Einsohn, A. (2006). *The Copyeditor's Handbook: A Guide for Book Publishing and Corporate Communications* (2nd ed.). Berkeley, CA: University of California.

Garner, B. A. (2003). *Garner's Modern American Usage* (2nd ed.). New York: Oxford University Press.

Merriam-Webster's Collegiate Dictionary (11th ed.). (2007). Springfield, MA: Merriam-Webster.

Microsoft Corporation Editorial Style Board. (2004). *Microsoft Manual of Style for Technical Publications* (3rd ed.). Redmond, WA: Microsoft Press.

Popham, W. J. (2010). *Classroom Assessment: What Teachers Need to Know* (6th ed.).

Upper Saddle River, NJ: Pearson. Russell, M. K., and Airasian, P. (2011).

Classroom Assessment (7th ed.). New York: McGraw-Hill.

Words into Type (3rd ed.). (1974). Upper Saddle River, NJ: Prentice-Hall.

---

[4] Hume, The Anatomy of Web Fonts

Appendix 2-D

Item Review Checklist

Item Review Checklist

Tier 1 – Sufficiency/Appropriateness of the Phenomenon to Assess the Performance Expectation

The elements in this tier are critical. If any of the following conditions are unmet, the item or cluster cannot move forward.

- ☐ Is the phenomenon based on a specific real-world scenario and focused enough to get the student to investigate what the Performance Expectation (PE) intends for them to investigate across **all three dimensions** (i.e., the students' application of the Practice in the context of the Disciplinary Core Idea [DCI] and Crosscutting Concepts [CCC] as intended by the PE is sufficient to make sense of the phenomena)?
- ☐ Is there an appropriate science-related activity that is puzzling and/or intriguing for students to engage in? Is the scenario focused on real-world observations that students can connect with or have direct experience with?
- ☐ Is the context and complexity of the phenomenon grade-appropriate?
- ☐ Cluster Task Statement: Does the "call to action" reflect the end goal of the interactions to be answered? Does the statement make sense? Is this an engaging and reasonable outcome to work towards?
- ☐ Is the phenomenon presented in way(s) that all students can access and comprehend it based on information provided (including text, graphics, data, images, animations, etc.)? Is the phenomenon free of cultural bias, insensitivity or depreciation of unsafe situations?

Tier 2 – Review of Specific Elements by Component

Stimulus

Reading Load/Readability/Style

- ☐ Is the reading load appropriate for the grade (i.e., the amount of text minimized to reduce cognitive load)?
- ☐ Is the language and vocabulary appropriate for the grade?
- ☐ Non-specific vocabulary should be one grade level lower than the tested grade.
- ☐ Science vocabulary should be part of the "Science Vocabulary Students Are Expected to Know" in the item specifications.
- ☐ Is all of the information in the stimulus necessary for the student to complete the item interactions?
- ☐ Is language consistent throughout the cluster (i.e., does not switch between steam and vapor)?

- ☐ Is everything in the active voice (i.e., avoids unnecessary and unclear passive construction)?

Measurement/Units

- ☐ Are the data in SI units? Check style guide for exceptions.
- ☐ Are units of measurement introduced or defined before they are used in graphs/tables?
- ☐ Are the dependent/independent variables on the correct axes or in the correct columns?
- ☐ Are the graphs/tables/pictures free of extraneous information and appropriate for the grade level?
- ☐ Is there information included in graphs/pictures/tables that is not necessary and can be removed?
- ☐ Do the graphs/tables/pictures depend on color? Is there another way to represent the difference in the data other than by color (e.g., using patterns)?

Data Source and Scientific Reference

- ☐ Is content both accurate and appropriate in its context?
- ☐ Are the data sources appropriate for the subject/grade and taken from reliable academic sources?
- ☐ Does the item use the most up-to-date explanation?

Formatting

- ☐ Is everything presented within the browser dimensions (1024x768) without horizontal scrolling?
- ☐ Are the tables/graphs/etc. laid out in a way that is easy to read?
- ☐ Are details and text in animations easy to see? Are labels in diagrams easy to read?
- ☐ Is the average file size appropriate for test delivery (approximately 100KB, 250KB maximum)?

Item

Interaction and Alignment to Specifications

- ☐ Does the item make sense if you are responding to the interactions as if you are the student in the intended grade-level?
- ☐ Does the interaction require the student to demonstrate the science practice and/or content that the PE is assessing them on?
- ☐ Are the interactions grade level/developmentally appropriate and do they follow a logical progression? Do the interactions use appropriate scaffolding to guide students in making sense of the phenomena?
- ☐ Do the interactions align with the task demands?

- Do the interactions avoid redundancy? Do the student interactions follow a coherent progression?
- Do the student interactions follow a coherent progression? Does the order of the interactions allow students to make sense of the phenomenon or problem?
- Is the item stem worded in a way that makes the intent of the interaction clear to the student?
- Is it clear to the student what they will be scored on in the interaction?
- Is the language (e.g., words, phrases) consistent throughout the stimulus and items?

Grade Appropriate

- Is the content within the item accurate and grade appropriate?
- Are the correct units used? Are the units grade appropriate? Where necessary, are the abbreviations of the units introduced?
- Is the number of item parts/scoring assertions appropriate for the grade level?
- Is the mathematics level appropriate for the grade being tested?

Formatting

- Is everything presented within the browser frame without horizontal scrolling?
- Are the tables/graphs/etc. easy to read? Are the images created in an appropriate color palette per the Style Guide?
- Are details and text in animations easy to see?

Tier 3 – Review of the Scoring and Assertion(s)

Scoring Accuracy

- Do the interactions/task provide clear guidance on how student responses will be scored/interpreted?
- Are scores assigned appropriately as correct or incorrect?
- Are the dependencies logical?
- Are any of the scoring assertions exclusive (i.e., the student can get only one assertion correct and not another at any given time)?
- Is the correct answer clear and distinct from the distractors?
- Does the scoring result in an appropriate distribution of points?

<u>Scoring Assertions</u>

- ☐ Is the appropriate wording used for each scoring assertion (e.g., <What the student did as a response> provides evidence of an understanding of/ability to <inference about student's ability relative to the PE being measured>)?
- ☐ Does the inference follow from the data?
- ☐ Are the assertions specific to the individual interactions (i.e., does not just repeat the PE)?
- ☐ Are the scoring assertions in the same order as the interactions?
- ☐ Does the wording of the scoring assertion make it very clear which interaction and action it refers to?

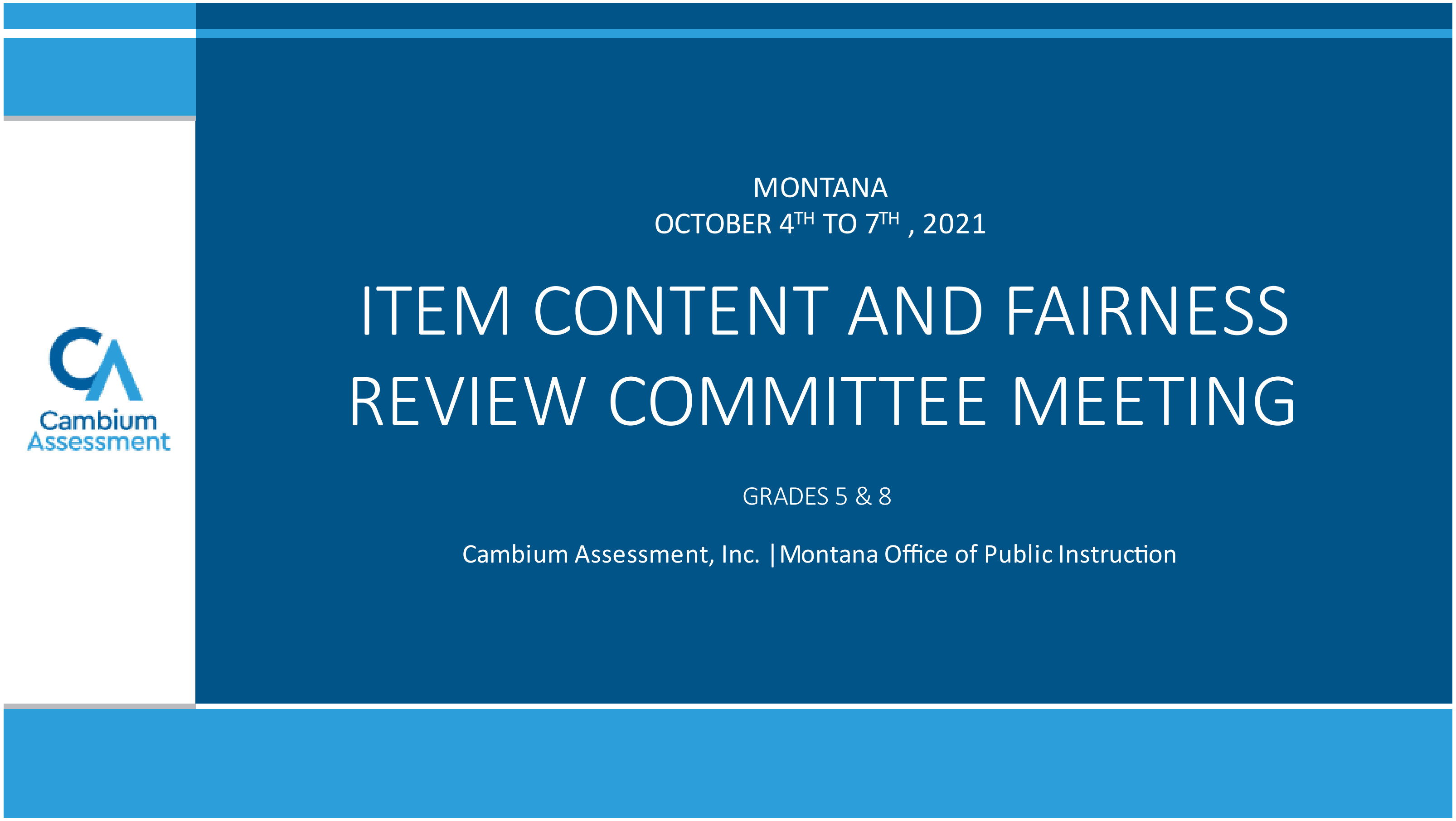
Strategies for Editing Text to Produce Plain Language

- Reduce excessive length
- Use common words
- Avoid ambiguous words
- Limit irregularly spelled words
- Avoid inconsistent naming and graphic conventions
- Avoid multiple terms for the same concept
- Limit the use of embedded clauses and phrases
- Avoid the passive voice

Appendix 2-E

Content Advisory Committee (CAC) Review Training Slides

Content Advisory Committee Review Training Slides



Welcome

- Montana OPI
- CAI
  - » Hibbah Haddam (ES Room)
  - » Mark Werner (MS Room)

Test Security

- Non-disclosure agreements should have already been signed.
- All test materials viewed during this meeting are considered secure.
- Do NOT discuss test material content outside of this meeting.

3

Agenda Day 1

- Hour 1: Introductions and Training; logging into content rater
- Hour 2-3: Individual review of items in content rater and group discussion

4

Overview of Item Review Training

- Steps in the Development Process
- Describe Structure of Three -Dimensional Standard, Clusters & Standalones
- Describe Scoring Assertions
- Role of the ContentReview Committee
- Content Review Process
- Norms & Participant Guidelines

5

Steps in the Development Process

Steps in Development Process

- Cambium Writes Clusters & Standalones
- Cambium Internal Review (Content & Fairness)
- OPI Review (Content & Fairness)
- **Educator & Stakeholder Review (Content & Fairness)**
- Field Test with Students
  - Rubric Validation and Data Review
- Operational Use

7

Structure of Three Dimensional Standards, Clusters, and Standalones

Next Generation Science Standards (NGSS)

- New science assessment looks very different
  - Focus on clusters
    - Aligned to a single performance expectation
    - Consist of multiple interactions

9

Three Dimensional Standards (NGSS)

Students who demonstrate understanding can:

MS-LS2-4. **Construct an argument supported by empirical evidence that changes to physical or biological components of an ecosystem affect populations.** [Clarification Statement: Emphasis is on recognizing patterns in data and making warranted inferences about changes in populations, and on evaluating empirical evidence supporting arguments about changes to ecosystems.]

The performance expectation above was developed using the following elements from the NRC document *A Framework for K-12 Science Education*:

Science and Engineering Practices	Disciplinary Core Ideas	Crosscutting Concepts
Engaging in Argument from Evidence Engaging in argument from evidence in 6–8 builds on K–5 experiences and progresses to constructing a convincing argument that supports or refutes claims for either explanations or solutions about the natural and designed world(s). • Construct an oral and written argument supported by empirical evidence and scientific reasoning to support or refute an explanation or a model for a phenomenon or a solution to a problem. - - - - - - - - - - - - - - - - ***Connections to Nature of Science*** **Scientific Knowledge is Based on Empirical Evidence** • Science disciplines share common rules of obtaining and evaluating empirical evidence.	**LS2.C: Ecosystem Dynamics, Functioning, and Resilience** • Ecosystems are dynamic in nature; their characteristics can vary over time. Disruptions to any physical or biological component of an ecosystem can lead to shifts in all its populations.	**Stability and Change** • Small changes in one part of a system might cause large changes in another part.

Connections to other DCIs in this grade-band:
MS.LS4.C ; **MS.LS4.D** ; **MS.ESS2.A** ; **MS.ESS3.A** ; **MS.ESS3.C**

Articulation of DCIs across grade-bands:
3.LS2.C ; **3.LS4.D** ; **HS.LS2.C** ; **HS.LS4.C** ; **HS.LS4.D** ; **HS.ESS2.E** ; **HS.ESS3.B** ; **HS.ESS3.C**

Common Core State Standards Connections:

ELA/Literacy -

RST.6-8.1 Cite specific textual evidence to support analysis of science and technical texts. (MS-LS2-4)

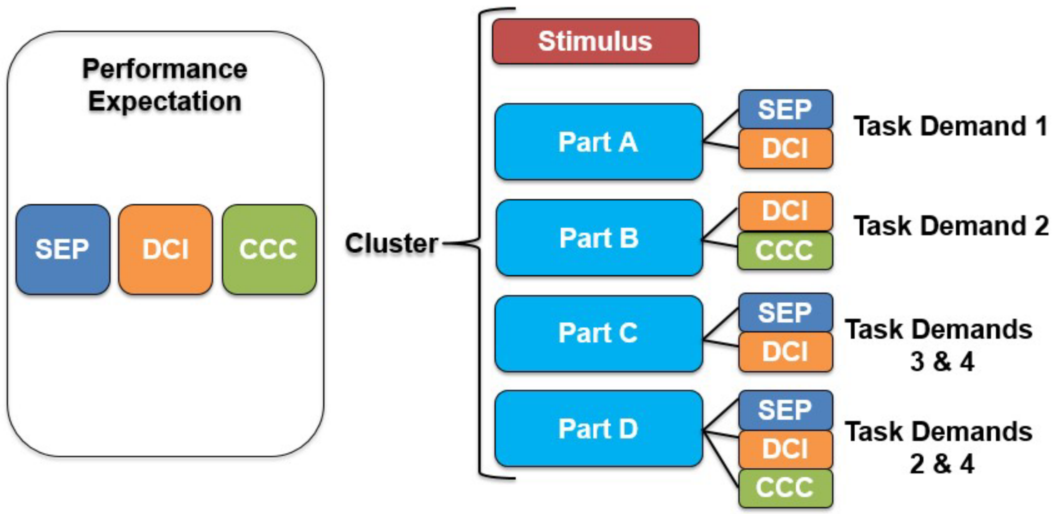
RI.8.8 Trace and evaluate the argument and specific claims in a text, assessing whether the reasoning is sound and the evidence is relevant and sufficient to support the claims. (MS-LS2-4)

WHST.6-8.1 Write arguments to support claims with clear reasons and relevant evidence. (MS-LS2-4)

WHST.6-8.9 Draw evidence from literary or informational texts to support analysis, reflection, and research. (MS-LS2-4)

10

Structure of Cambium Clusters



11

Structure of Cambium Clusters

Task Demands
1. Based on the provided data, identify, describe or illustrate a claim regarding the relationship between a characteristic animal behavior and/or specialized plant structure and the probability of successful reproduction in the species.
2. Identify, summarize, or organize given data or other information to support or refute a claim regarding the relationship between a characteristic animal behavior and/or specialized plant structure and the probability of successful reproduction in the species.
3. Sort inferences about the relationship of behaviors or structures to breeding success into those that are supported by the data, contradicted by the data, or neither, or some similar classification.
4. Select supporting evidence from competing sources based on the reliability of statistical relationships, how representative the sample is, or study design.
5. Construct an argument using scientific reasoning drawing on credible evidence to explain the relationships of animal behaviors or plant structures to reproductive success. (Hand scored CR)
6. Identify additional evidence that would help clarify, support, or contradict a hypothesized relationship or causal argument.
7. Identify or describe alternate explanations and the data needed to distinguish among them.

12

Willow populations in Yellowstone National Park have increased since wolves were reintroduced to the park in 1995.

Willows are small trees that grow best in marshlike environments. After studying the Yellowstone food web shown in Diagram 1 and the population data for the park shown in Table 1, students arrive at two different hypotheses.

Diagram 1. Yellowstone Food Web

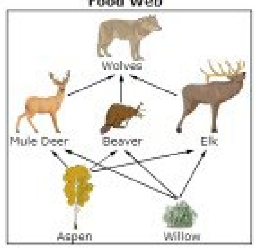


Table 1. Yellowstone Population Data

	Wolves	Elk	Beaver	Mule Deer
1995	31	16,791	10	2,014
2004	171	8,335	120	2,014

Note: These data are approximate.

Hypothesis 1:
When wolves were reintroduced to Yellowstone, the wolves preyed upon the elk, which allowed the beavers to eat more willow. This led to more beavers and beaver dams. Beaver dams create marsh environments that willows do well in, allowing the willow's population to increase.

Part A

Click on each box and select a word/phrase that completes the table with the Yellowstone population data from 1995 and 2004 and the hypothesis those data support.

Table 2. Summary of Yellowstone Population Data and Supported Hypotheses

Data	Hypothesis Supported
Elk population [▾]	[▾]
Beaver population [▾]	[▾]
Mule deer population [▾]	[▾]

Part B

Which hypothesis is best supported by the evidence?

- Ⓐ All of the evidence is consistent with Hypothesis 1.
- Ⓑ All of the evidence is consistent with Hypothesis 2.
- Ⓒ Most of the evidence is consistent with Hypothesis 1.
- Ⓓ Most of the evidence is consistent with Hypothesis 2.
- Ⓔ The evidence does not favor either hypothesis.

Part C

Aspen trees are shown in Diagram 1. Moose and bison are two plant-eating animal species that are not shown in Diagram 1 but are also part of the Yellowstone food web.

Based on Hypothesis 2, click on each box to select a word/phrase to make a prediction about what would happen to the moose, bison, and aspen tree populations after the reintroduction of wolves.

Table 3. Population Predictions

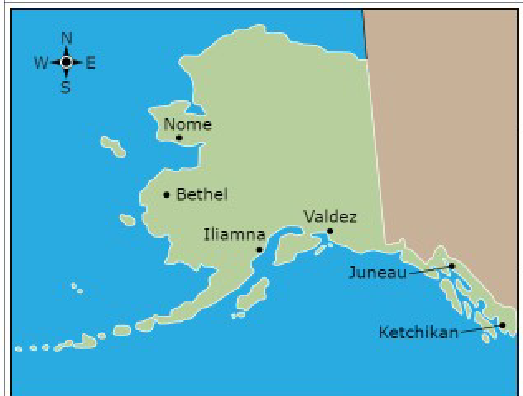
Species	Population after Wolf Reintroduction	Reason for Impact on Population
Moose	[▾]	[▾]
Bison	[▾]	[▾]
Aspen tree	[▾]	[▾]

13

Standalone

Table 1 shows six locations in Alaska. It also shows how much rain three of the locations received in March.

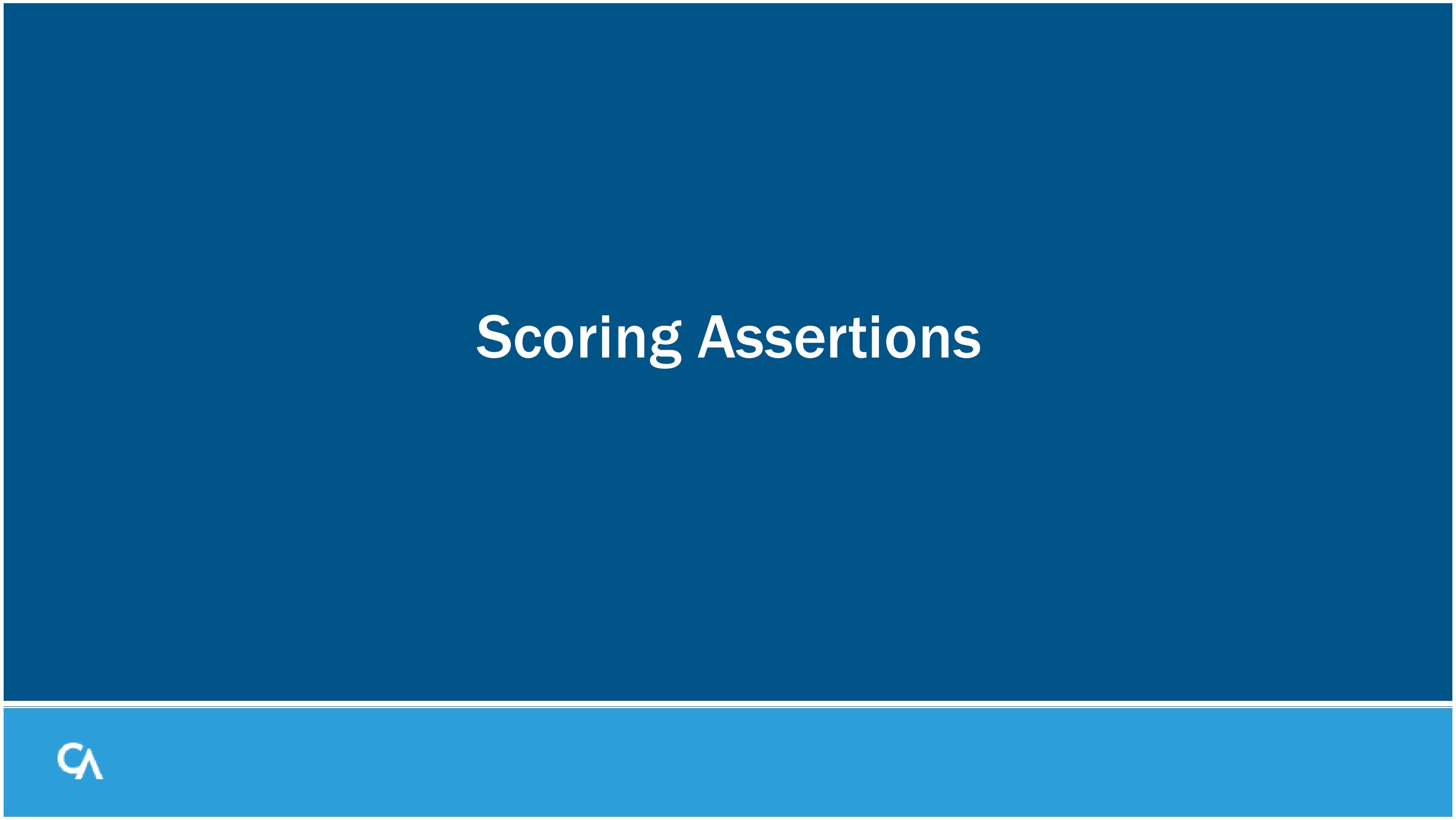
Table 1. Average Rainfall Amounts in Alaska during March

Map of Alaska	Location	Average Rainfall (centimeters)
	Ketchikan	27.7
	Valdez	11.5
	Nome	1.7

Based on the data in Table 1, rank the other three locations by the average rainfall they received in March. Click on the blank boxes to make your selections.

Most Rain	
↓	
Least Rain	

14


Scoring Assertions

Scoring Assertions

- Within each item are a series of explicit assertions that can be made about the knowledge and skills that a student has demonstrated based on specific features of the student's responses
- Scoring assertions tell us:
  - What the student did
  - What inference can be made from that action
- e.g.
  - The student correctly graphs data points, providing some evidence of the ability to construct a graph showing the relationship between two variables.

16

Scoring Assertions

Score Result

Your response earned 0 point(s) of a possible 15

Scoring Criteria	Your answer
The student identified that the elk population decreased between 1995 and 2004 giving evidence of understanding of how to interpret the data presented in Table 1.	✖
The student identified that the elk population decreased between 1995 and 2004 and that data supports both hypotheses giving evidence of understanding of how to use evidence to support an argument.	✖
The student identified that the beaver population increased between 1995 and 2004 giving evidence of understanding of how to interpret the data presented in Table 1.	✖
The student identified that the beaver population increased between 1995 and 2004 and that data supports Hypothesis 1 giving evidence of understanding of how to use evidence to support an argument.	✖
The student identified that the mule deer population had no change between 1995 and 2004 giving evidence of understanding of how to interpret the data presented in Table 1.	✖
The student identified that the mule deer population had no change between 1995 and 2004 and that data supports neither hypothesis giving evidence of understanding of how to use evidence to support an argument.	✖
The student identified that most of the evidence is consistent with Hypothesis 1 or the student's conclusion matches their correct inferences, given their characterization of the data, providing some evidence of his or her ability to summarize information and draw a conclusion.	✖
The student predicted that the moose population would decrease giving evidence of the ability to extrapolate information from a given food web.	✖
The student predicted that the moose population would decrease because they would be preyed on by wolves giving evidence of the ability to support a correct prediction with logical reasoning.	✖
The student predicted that the bison population would decrease giving evidence of the ability to extrapolate information from a given food web.	✖
The student predicted that the bison population would decrease because they would be	

17

Role of Content Review Committee

Points of Content Review

1. Phenomenon
2. Language
3. Scientific Content
4. Interactions/Alignment
5. Scoring Assertions
6. Fairness

19

1. Phenomenon

- Is the phenomenon based on a specific real-world scenario and focused enough to get the student to investigate what the PE intends for them to investigate?
- Is there an appropriate science-related activity that is puzzling and/or intriguing for students to engage in?
- Does the Cluster Task Statement (CTS) reflect the end goal of the interactions to be answered?
- Is the phenomenon presented in way(s) that all students can access and comprehend it based on information provided (including text, graphics, data, images, animations, etc.)

20

2. Language

- Is the language and vocabulary appropriate for the grade?
- Is all of the information in the stimulus necessary for the student to complete the interactions?
- Is everything in active voice (i.e., avoids unnecessary and unclear passive construction)?

21

3. Scientific Content

- Is scientific content both accurate and appropriate in its context?
- Are the graphs/tables/pictures scientifically correct and appropriate for the grade level?
- Are the data in appropriate units? Are units of measure introduced or defined before they are used in graphs/tables?

22

4. Interactions/Alignment

- Is the item two or three dimensionally aligned?
- Do the interactions make sense if you are responding to the interactions as if you are the student in the intended grade-level?
- Do the interactions require the student to demonstrate the science practice and/or content that the PE is assessing them on?
- Are the interactions grade level/developmentally appropriate and do they follow a logical progression?

23

5. Scoring Assertions

- Do the interactions/task provide clear guidance on how student responses will be scored/interpreted?
- Are scores assigned appropriately as correct or incorrect? Are the dependencies logical?
- Is the correct answer clear and distinct from the distractors?
- Is the appropriate wording used for each scoring assertion?

24

6. Fairness

Purpose:

- Test materials need to be free from unnecessary barriers to the success of diverse groups of students.

Why?

1. Valid measurement of student achievement.
2. Positive student experience.

25

Assessment Fairness

Assessment content is…

1. free of bias and stereotypes.
2. sensitive to student and community beliefs and experiences.
3. accessible to all students, to the greatest extent possible.

Assessment Fairness

Assessment content is free of bias and stereotypes.

- Assessment content does not express bias or present stereotypes of people of different groups.
- Assessment content demonstrates respect for people of different groups.
- Assessment content is inclusive and reflects the diversity of the community.

Assessment Fairness

Assessment content is free of bias and stereotypes.

- Gender
- Race, Ethnicity, Culture
- Religion
- Age
- Disability
- Socioeconomic

Assessment Fairness

Assessment content is sensitive to student and community beliefs and experiences.

1. Assessment content does not include topics that are considered controversial by the community, unless these topics are part of the learning standards.

2. Assessment content avoids emotionally-charged topics.

3. Assessment content does not give advice or promote specific morals, unless these are universally accepted.

4. Assessment content does not depict people engaged in dangerous activities.

29

Assessment Fairness

Assessment content is accessible to all students to the greatest extent possible.

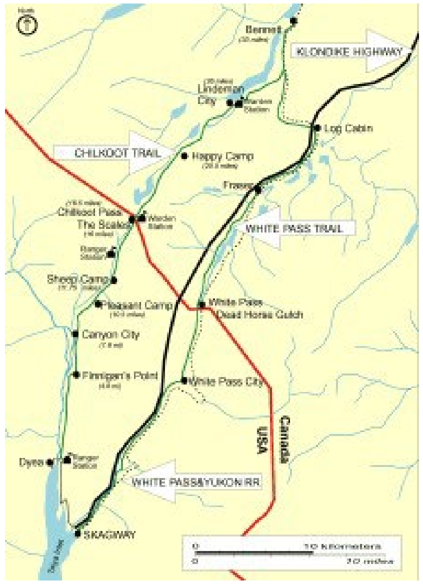
1. Language used in the assessment is understandable and inclusive.

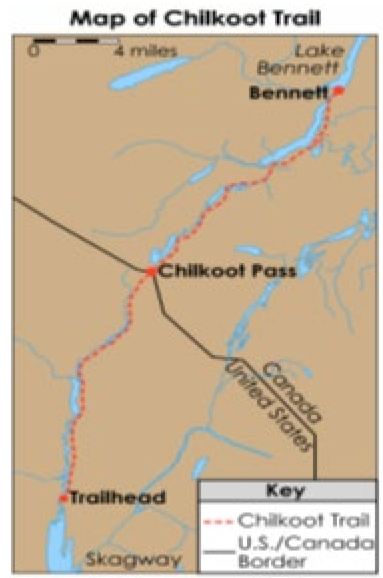
2. Assessment content does not rely on vocabulary or background knowledge that would be significantly more or less familiar to a student because of their demographic group. (Differential familiarity)

3. Accommodations and supports are available to students when needed.

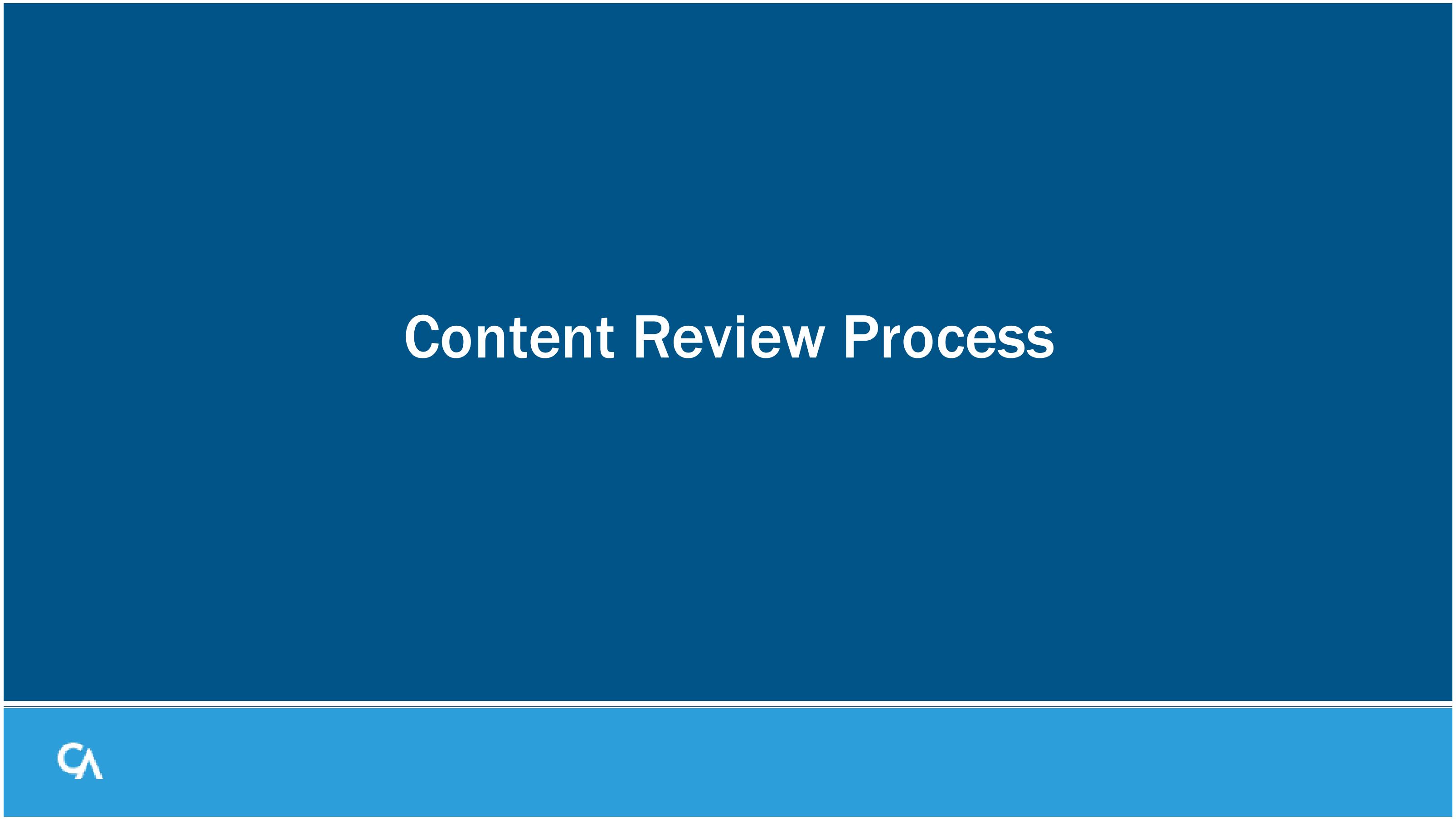
30

Assessment Fairness





31


Content Review Process

Content Review Process

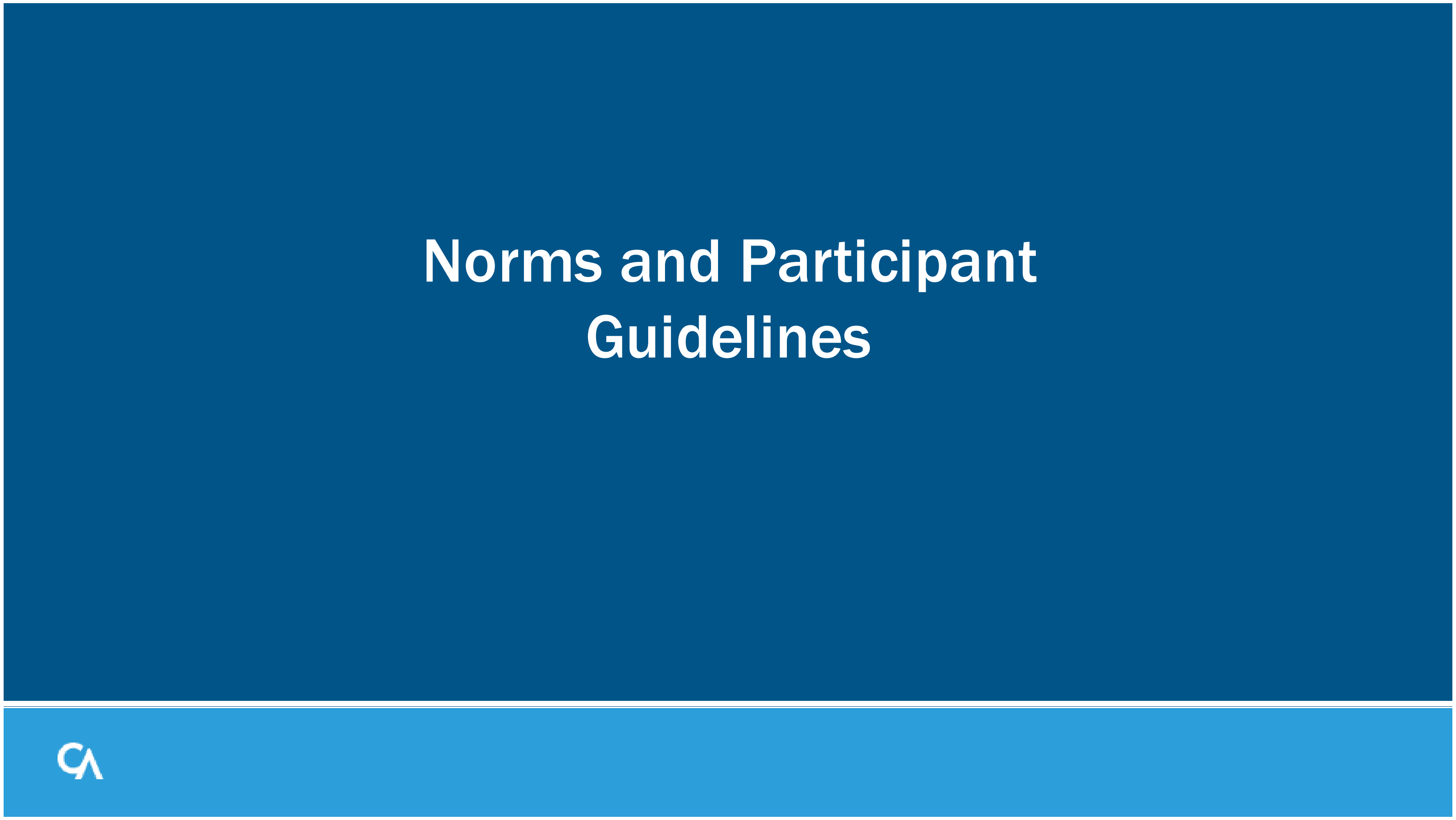
1. Enter Content Rater system and begin batch.
2. Review content of each item. Read through item and take it like the students would.
   a. Utilize standards if needed
   b. Document comments in content rater for discussion

33

Content Review Process

- Cambium facilitator will lead discussion of items.
  - Facilitator documents group decision/feedback from committee directly into ITS.
    - » Discuss any suggested revisions that will help improve the item
  - Cambium facilitator also documents review outcome recommendation in ITS.
    - » AAA-Accept as Appears
    - » AAR-Accept as Revised
    - » Recommend to Reject
- Later, feedback will be discussed with OPI

34


Norms and Participant
Guidelines

Participant Guidelines

- Keep phones turned off & stowed while completing review.
- Utilize Go-To meeting chat for any questions or concerns during the individual review time.
- Do not speak (or write) about specific item content outside of this review. Be aware of other people who may overhear you or see your screen. Item content MUST remain secure.
- If you have any questions about the review or procedures, feel free to reach out to Cambium staff or OPI via Go-To Meeting.

36

Norms

- Listen actively and attentively to each other and the facilitator.
- Take turns providing feedback, encourage everyone to participate.
- Ask for clarification when needed.
- Question the cluster, not the writer.
- Avoid off topic & side conversations.
- Build on one another’s comments; work towards a group consensus.

37


Questions?

Appendix 2-F

Content Advisory Committee Participant Details

Content Advisory Committee Participant Details

Table F-1. Content Advisory Committee Participants, Science

State/Item Bank	Date	Location	Number of Committee Members	CAC Committee Member Demographic Summary	Number of Items Reviewed	Number of Items Approved by Committees
Connecticut	February '17	Cromwell, CT	41	**Gender:** 78% Female, 22% Male	45	31
	May '17	New Britain, CT	42	**Gender:** 74% Female, 26% Male	40	38
	October '17	New Britain, CT	41	**Gender:** 80% Female, 20% Male	75	64
	November '17	New Britain, CT	35	**Gender:** 83% Female, 17% Male	41	32
	January/ February '18	New Britain, CT	33	**Gender:** 82% Female, 18% Male	42	25
	October '18	New Britain, CT	45	**Gender:** 84% Female, 16% Male	84	54
	November '18	New Britain, CT	49	**Gender:** 86% Female, 14% Male	235	200
	December '18	New Britain, CT	32	**Gender:** 81% Female, 19% Male	56	55
	January '19	New Britain, CT	44	**Gender:** 82% Female, 18% Male	65	59
	September '19	Rocky Hill, CT	50	**Gender:** 82% Female, 18% Male	60	57
	November '19	Cromwell, CT	44	**Gender:** 80% Female, 20% Male **Ethnicity:** 5% Hispanic or Latino, 93% White or Caucasian, 2% Preferred Not to Answer **Region:** 14% Rural, 59% Suburban, 16% Urban, 11% Not Applicable **Teaching Experience:** 2% None, 9% 1 to 5 years, 9% 6 to 10 years, 30% 11 to 15 years, 25% 16 to 20 years, 25% More than 20 years	171	153
	January '20	Cromwell, CT	57	**Gender:** 75% Female, 25% Male **Ethnicity:** 5% Black or African American, 2% Franco-American, 5% Hispanic or Latino, 88% White or Caucasian **Region:** 14% Rural, 63% Suburban, 19% Urban, 4% Not Applicable	190	161

State/Item Bank	Date	Location	Number of Committee Members	CAC Committee Member Demographic Summary	Number of Items Reviewed	Number of Items Approved by Committees
				Teaching Experience: 12% 1 to 5 years, 14% 6 to 10 years, 25% 11 to 15 years, 21% 16 to 20 years, 28% More than 20 years		
	July ‘20[a]	Virtual	23	**Gender:** 83% Female, 17% Male **Ethnicity:** 4% Black or African American, 91% White or Caucasian, 4% Prefer Not to Answer	48	44
	July ‘21[a]	Virtual	68	**State:** 22% Connecticut, 4% Hawaii, 9% Idaho, 4% Montana, 4% Oregon, 4% Rhode Island, 13% South Dakota, 4% Utah, 3% Vermont, 12% West Virginia, 13% Wyoming, 4% Did Not Respond **Teaching Experience:** 9% 1 to 5 years, 18% 6 to 10 years, 29% 11 to 15 years, 19% 16 to 20 years, 25% More than 20 years	26	26
	September ‘21	Virtual	27	**Gender:** 74% Female, 26% Male **Ethnicity:** 4% Black or African American, 96% White or Caucasian **Region:** 41% Suburban, 11% Urban, 7% Not Applicable, 41% Did Not Respond **Teaching Experience:** 7% 1 to 5 years, 15% 6 to 10 years, 15% 11 to 15 years, 15% 16 to 20 years, 48% More than 20 years	149	120
	June/ August ‘22[a]	Virtual	26	**State:** 8% Connecticut, 11% Idaho, 4% North Dakota, 4% Oregon, 19% Rhode Island, 27% Utah, 8% West Virginia, 19% Wyoming **Gender:** 77% Female, 23% Male **Ethnicity:** 15% Hispanic or Latino, 81% White or Caucasian, 4% Race/Ethnicity Not Listed **Region:** 19% Rural, 23% Urban, 31% Suburban, 27% Did Not Respond **Teaching Experience:** 15% 11 to 15 years, 23% 6 to 10 years, 35% 16 to 20 years, 27% More than 20 years	65	63
	July ‘22	Virtual	21	**Gender:** 76% Female, 24% Male **Ethnicity:** 10% Hispanic or Latino, 86% White or Caucasian, 4% Multiracial or Biracial **Region:** 76% Suburban, 10% Urban, 14% Did Not Respond **Teaching Experience:** 19% 6 to 10 years, 4% 11 to 15 years, 10% 16 to 20 years, 67% More than 20 years	62	56

State/Item Bank	Date	Location	Number of Committee Members	CAC Committee Member Demographic Summary	Number of Items Reviewed	Number of Items Approved by Committees
Hawaii	July ‘17	Honolulu, HI	22	**Gender:** 64% Female, 36% Male **Ethnicity:** 5% Black or African American, 5% Chinese and White, 9% Filipino, 14% Hawaiian, 9% Hispanic or Latino, 14% Japanese, 41% White or Caucasian, 5% Did Not Respond **Teaching Experience:** 64% General Education, 5% General Education with SPED Certification, 5% SPED Teacher, 23% Other, 5% Did Not Respond	25	[b]
	September ‘17	Honolulu, HI	20	**Gender:** 75% Female, 25% Male **Ethnicity:** 5% Black or African American, 10% Filipino, 10% Hispanic or Latino, 15% Japanese, 50% White or Caucasian, 10% Did Not Respond **Teaching Experience:** 65% General Education, 15% General Education with SPED Certification, 20% Other	65	[b]
	October ‘18	Honolulu, HI	28	**Gender:** 83% Female, 17% Male **Ethnicity:** 31% Asian, 7% Asian Pacific Islander, 3% Hawaiian, 10% Hispanic or Latino, 28% White or Caucasian, 10% Two or More, 10% Did Not Respond **Teaching Experience:** 83% General Education, 24% Other	85	79
	February/ March ‘19	Honolulu, HI	21	**Gender:** 80% Female, 20% Male **Ethnicity:** 50% Asian, 35% White or Caucasian, 15% Two or More **Teaching Experience:** 65% General Education, 5% General Education with SPED Certification, 5% SPED Teacher, 25% Other	44	44
	June/ July ‘20	Virtual	17	**Gender:** 18% Female, 12% Male, 70% Did Not Respond **Ethnicity**: 6% Asian or Pacific Islander, 6% Multiracial or Biracial, 18% White or Caucasian, 70% Did Not Respond **Region:** 12% Rural, 12% Suburban, 76% Did Not Respond **Teaching Experience:** 6% 6 to 10 years, 12% 11 to 15 years, 12% More than 20 years, 70% Did Not Respond	344	324
	July ‘20[a]	Virtual	28	**State:** 14% Connecticut, 4% Hawaii, 14% Idaho, 14% Montana, 7% Oregon, 4% Rhode Island, 4% Utah, 7% Vermont, 11% West Virginia, 7% Wyoming, 14% Did Not Respond **Gender:** 86% Female, 14% Male	90	90

State/Item Bank	Date	Location	Number of Committee Members	CAC Committee Member Demographic Summary	Number of Items Reviewed	Number of Items Approved by Committees
				Ethnicity: 4% Black or African American, 46% White or Caucasian, 50% Did Not Respond **Region:** 7% Rural, 14% Suburban, 14% Urban, 64% Did Not Respond **Teaching Experience:** 3% 6 to 10 years, 11% 11 to 15 years, 14% 16 to 20 years, 18% More than 20 years, 54% Did Not Respond		
	June/ August '22[a]	Virtual	25	**State:** 12% Connecticut, 16% Idaho, 16% North Dakota, 16% Oregon, 20% Rhode Island, 12% Utah, 4% West Virginia, 4% Wyoming **Gender:** 88% Female, 12% Male **Ethnicity:** 8% Hispanic or Latino, 4% Multiracial or Biracial, 80% White or Caucasian, 8% Did Not Respond **Region:** 20% Rural, 16% Suburban, 12% Urban, 52% Did Not Respond **Teaching Experience:** 4% 1 to 5 Years, 32% 6 to 10 Years, 20% 11 to 15 Years, 20% 16 to 20 Years, 16% More than 20 Years, 8% Did Not Respond	46	46
	July '22	Honolulu, HI	9	**Gender:** 67% Female, 22% Male, 11% Non-Binary **Ethnicity:** 44% Asian/Pacific Islander, 11% Hispanic or Latino, 11% Multiracial or Biracial, 22% White or Caucasian, 11% Did Not Respond **Region:** 11% Hilo, 11% Maui, 78% Oahu **Teaching Experience:** 11% 1 to 5 years, 11% 6 to 10 years, 11% 11 to 15 years, 22% 16 to 20 years, 44% More than 20 years	45	44
	July '22	Honolulu, HI	9	**Gender:** 67% Female, 22% Male, 11% Non-Binary **Ethnicity:** 44% Asian/Pacific Islander, 11% Hispanic or Latino, 11% Multiracial or Biracial, 22% White or Caucasian, 11% Did Not Respond **Region:** 11% Hilo, 11% Maui, 78% Oahu **Teaching Experience:** 11% 1 to 5 years, 11% 6 to 10 years, 11% 11 to 15 years, 22% 16 to 20 years, 44% More than 20 years	306	306
ICCR	March '18	Virtual	38	**State:** 45% Connecticut, 5% Hawaii, 3% Indiana, 3% Maryland, 8% Oregon, 8% Utah, 26% West Virginia, 3% Wyoming **Gender:** 74% Female, 26% Male	152	[b]

State/Item Bank	Date	Location	Number of Committee Members	CAC Committee Member Demographic Summary	Number of Items Reviewed	Number of Items Approved by Committees
	July ‘20[a]	Virtual	6	**State:** 17% Connecticut, 17% Idaho, 17% Oregon, 17% Rhode Island, 33% Did Not Respond **Gender:** 83% Female, 17% Male **Ethnicity**: 33% White or Caucasian, 67% Did Not Respond **Region:** 17% Suburban, 83% Did Not Respond **Teaching Experience:** 33% 16 to 20 years, 67% Did Not Respond	57	56
	July ‘21[a]	Virtual	68	**State:** 22% Connecticut, 4% Hawaii, 9% Idaho, 4% Montana, 4% Oregon, 4% Rhode Island, 13% South Dakota, 4% Utah, 3% Vermont, 12% West Virginia, 13% Wyoming, 4% Did Not Respond **Teaching Experience:** 9% 1 to 5 years, 18% 6 to 10 years, 29% 11 to 15 years, 19% 16 to 20 years, 25% More than 20 years	141	141
	July ‘21	Virtual	45	**State:** 33% Connecticut, 9% Hawaii, 4% Idaho, 2% Montana, 7% North Dakota, 4% Oregon, 18% South Dakota, 2% Vermont, 4% West Virginia, 13% Wyoming **Gender:** 80% Female, 18% Male, 2% Did Not Respond **Ethnicity:** 4% Asian or Pacific Islander, 2% Black or African American, 4% Hispanic or Latino, 87% White or Caucasian, 2% Did Not Respond **Region:** 36% Rural, 24% Suburban, 20% Urban, 20% Did Not Respond **Teaching Experience:** 2% None, 2% Less than 1 year, 11% 1 to 5 years, 33% 6 to 10 years, 16% 11 to 15 years, 9% 16 to 20 years, 24% More than 20 years, 2% Did Not Respond	163	158
	June/ August ‘22[a]	Virtual	12	**State:** 17% Connecticut, 17% Idaho, 58% North Dakota, 8% Oregon **Gender:** 92% Female, 8% Male **Ethnicity:** 67% White or Caucasian, 33% Did Not Respond **Region:** 33% Suburban, 33% Urban, 33% Did Not Respond **Teaching Experience:** 17% 6 to 10 Years, 42% 11 to 15 Years, 8% More than 20 Years, 33% Did Not Respond	121	118
Idaho	December ‘18	Boise, ID	21	Not Collected	241	230
	October ‘19	Boise, ID	18	**Gender:** 83% Female, 11% Male, 6% Did Not Respond **Ethnicity:** 100% White or Caucasian	231	211

State/Item Bank	Date	Location	Number of Committee Members	CAC Committee Member Demographic Summary	Number of Items Reviewed	Number of Items Approved by Committees
				Region: 50% Rural, 17% Suburban, 22% Urban, 11% Not Applicable **Teaching Experience:** 11% 1 to 5 years, 22% 6 to 10 years, 17% 11 to 15 years, 11% 16 to 20 years, 28% 21 or more years, 11% Did Not Respond		
	July ‘20[a]	Virtual	2	**State:** 100% Hawaii **Gender:** 100% Female	12	12
	October ‘20	Virtual	[c]	Not Collected	14	14
	July ‘21[a]	Virtual	8	**Gender:** 88% Female, 13% Male **Ethnicity:** 100% White or Caucasian **Region:** 25% Rural, 25% Suburban, 50% Did Not Respond **Teaching Experience:** 38% 6 to 10 years, 38% 11 to 15 years, 13% 16 to 20 years, 13% More than 20 years	[b]	[b]
	November ‘21	Virtual	11	**Gender:** 91% Female, 9% Male **Ethnicity**: 73% White or Caucasian, 17% Did Not Respond **Region:** 18% Rural, 18% Suburban, 64% Did Not Respond **Teaching Experience:** 9% 1 to 5 years, 46% 6 to 10 years, 9% 11 to 15 years, 9% More than 20 years, 27% Did Not Respond	317	286
	June/ August ‘22[a]	Virtual	14	**State:** 29% Connecticut, 14% Oregon, 7% Rhode Island, 7% North Dakota, 21% Utah, 7% West Virginia, 14% Wyoming **Gender:** 71% Female, 29% Male **Ethnicity:** 7% Hispanic or Latino, 93% White or Caucasian, **Region:** 36% Rural, 21% Suburban, 14% Urban, 29% Did Not Respond **Teaching Experience:** 29% 6 to 10 Years, 21% 11 to 15 Years, 29% 16 to 20 Years, 21% More than 20 Years	12	12
	July ‘22	Virtual	5	**Gender**: 100% Female **Ethnicity**: 80% White or Caucasian, 20% Did Not Respond **Region:** 20% Rural, 40% Suburban, 40% Did Not Respond **Teaching Experience**: 40% 1 to 5 years, 40% 6 to 10 years, 20% 11 to 15 years	244	204
Montana	January ‘20	Helena, MT	15	Not Collected	149	139
	July ‘20[a]	Virtual	4	**State:** 25% Hawaii, 25% Idaho, 25% Oregon, 25% Rhode Island **Gender:** 75% Female, 25% Male	9	9

State/Item Bank	Date	Location	Number of Committee Members	CAC Committee Member Demographic Summary	Number of Items Reviewed	Number of Items Approved by Committees
				Ethnicity: 50% White or Caucasian, 50% Did Not Respond **Region:** 50% Urban, 50% Did Not Respond **Teaching Experience:** 50% More than 20 years, 50% Did Not Respond		
	October '20	Virtual	8	**Gender:** 13% Female, 88% Did Not Respond **Ethnicity**:13% White or Caucasian, 88% Did Not Respond **Region:** 13% Rural, 88% Did Not Respond **Teaching Experience:** 13% 16 to 20 years, 88% Did Not Respond	156	140
	July '21[a]	Virtual	68	**State:** 22% Connecticut, 4% Hawaii, 9% Idaho, 4% Montana, 4% Oregon, 4% Rhode Island, 13% South Dakota, 4% Utah, 3% Vermont, 12% West Virginia, 13% Wyoming, 4% Did Not Respond **Teaching Experience:** 9% 1 to 5 years, 18% 6 to 10 years, 29% 11 to 15 years, 19% 16 to 20 years, 25% More than 20 years	36	36
	October '21	Virtual	6	**Gender:** 83% Female, 17% Did Not Respond **Ethnicity**: 17% Hispanic or Latino, 67% White or Caucasian, 17% Did Not Respond **Region:** 67% Rural, 33% Did Not Respond **Teaching Experience:** 33% 6 to 10 years, 33% 11 to 15 years, 17% More than 20 years, 17% Did Not Respond	41	39
	June/ August '22[a]	Virtual	9	**State:** 22% Connecticut, 22% Idaho, 11% Oregon, 11% Rhode Island, 11% West Virginia, 22% Wyoming **Gender:** 78% Female, 22% Male **Ethnicity:** 100% White or Caucasian **Region:** 33% Rural, 22% Suburban, 11% Urban, 33% Did Not Respond **Teaching Experience:** 22% 6 to 10 Years, 33% 11 to 15 Years, 11% 16 to 20 Years, 33% More than 20 Years	13	13
Multi-State Science Assessment (Rhode Island and Vermont)	January '18	Providence, RI	42	**State:** 90% Rhode Island, 10% Vermont **Teaching Experience:** 69% General Education, 2% Bilingual Education, 14% Science Coordinator, 14% Other	73	58
	March '18	Providence, RI	34	**State:** 25% Rhode Island, 75% Vermont	107	90
	January '19	Concord, NH	21	**Gender:** 74% Female, 26% Male **Teaching Experience:** 69% General Education, 3% Special Education, 29% Other, 6% Not Applicable	116	97

State/Item Bank	Date	Location	Number of Committee Members	CAC Committee Member Demographic Summary	Number of Items Reviewed	Number of Items Approved by Committees
	November ‘19	Fairlee, VT	17	**State:** 29% Rhode Island, 6% Vermont, 65% Did Not Respond **Gender:** 23% Female, 12% Male, 65% Did Not Respond **Ethnicity**: 35% White or Caucasian, 65% Did Not Respond **Region:** 6% Rural, 17% Suburban, 77% Did Not Respond **Teaching Experience:** 6% 11 to 15 years, 17% 16 to 20 years, 12% More than 20 years, 65% Did Not Respond	136	118
	July ‘20[a]	Virtual	8	**State:** 13% Connecticut, 13% Hawaii, 13% Montana, 13% Oregon, 25% Oregon, 25% Did Not Respond **Gender:** 88% Female, 13% Male **Ethnicity**: 38% White or Caucasian, 63% Did Not Respond **Region:** 13% Suburban, 88% Did Not Respond **Teaching Experience:** 13% 6 to 10 years, 13% 11 to 15 years, 13% More than 20 years, 63% Did Not Respond	27	27
	July ‘21[a]	Virtual	68	**State:** 22% Connecticut, 4% Hawaii, 9% Idaho, 4% Montana, 4% Oregon, 4% Rhode Island, 13% South Dakota, 4% Utah, 3% Vermont, 12% West Virginia, 13% Wyoming, 4% Did Not Respond **Teaching Experience:** 9% 1 to 5 years, 18% 6 to 10 years, 29% 11 to 15 years, 19% 16 to 20 years, 25% More than 20 years	32	31
	August ‘21	Virtual	11	**State:** 45% Rhode Island, 55% Vermont **Gender:** 73% Female, 27% Male **Ethnicity:** 100% White or Caucasian **Region:** 36% Rural, 18% Suburban, 9% Urban, 36% Did Not Respond **Teaching Experience:** 27% 6 to 10 years, 18% 11 to 15 years, 27% 16 to 20 years, 27% More than 20 years	93	91
Oregon	August ‘17	Salem, OR	10	**Gender:** 90% Female, 10% Male **Region:** 50% Rural, 50% Urban **Teaching Experience:** 100% General Education, 10% Bilingual Education, 10% Special Education, 20% Administration	235	142
	August ‘18	Salem, OR	20	**Gender:** 80% Female, 20% Male **Ethnicity:** 95% White or Caucasian, 5% Other **Region:** 44% Rural, 56% Urban	257	200

State/Item Bank	Date	Location	Number of Committee Members	CAC Committee Member Demographic Summary	Number of Items Reviewed	Number of Items Approved by Committees
				Teaching Experience: 65% Bilingual Education, 65% Special Education, 55% Other		
	October ‘18	Salem, OR	11	**Gender:** 100% Female **Ethnicity:** 91% White or Caucasian, 9% Other **Region:** 45% Rural, 55% Urban **Teaching Experience:** 18% General Education, 91% Bilingual Education, 45% Special Education, 55% Other	60	30
	December ‘18	Virtual	16	**Gender:** 63% Female, 38% Male **Ethnicity:** 6% Asian, 94% White or Caucasian **Region:** 50% Suburban, 50% Urban **Teaching Experience:** 38% General Education, 63% Bilingual Education, 25% Special Education	62	48
	October ‘19	Salem, OR	17	**Gender:** 76% Female, 24% Male **Ethnicity:** 6% Asian, 88% White or Caucasian, 6% Other **Region:** 29% Rural, 71% Urban **Teaching Experience:** 82% General Education, 29% Bilingual Education, 18% Special Education	255	221
	July ‘20[a]	Virtual	9	**State:** 22% Idaho, 11% Vermont, 22% West Virginia, 11% Wyoming, 33% Did Not Respond **Gender:** 78% Female, 22% Male **Ethnicity**: 44% White or Caucasian, 56% Did Not Respond **Region:** 11% Rural, 11% Suburban, 78% Did Not Respond **Teaching Experience:** 11% 11 to 15 years, 11% 16 to 20 years, 22% More than 20 years, 56% Did Not Respond	22	20
	August ‘20	Virtual	21	**Gender:** 71% Female, 29% Male **Ethnicity:** 90% White or Caucasian, 5% Hispanic or Latino, 5% Native American **Region:** 5% Urban, 43% Suburban, 52% Rural **Teaching Experience:** 86% General Education, 81% Bilingual Education, 81% Special Education, 14% Administration, 5% Other	159	134
	August ‘21	Virtual	14	**Gender:** 86% Female, 14% Male **Ethnicity:** 86% White or Caucasian, 7% Asian and/or Pacific Islander, 7% Hispanic or Latino **Region:** 14% Urban, 72% Suburban, 14% Rural **Teaching Experience:** 64% General Education, 7% Bilingual Education, 7% Special Education, 22% Other	375	308

State/Item Bank	Date	Location	Number of Committee Members	CAC Committee Member Demographic Summary	Number of Items Reviewed	Number of Items Approved by Committees
	July ‘22	Virtual	14	**Gender:** 64% Female. 36% Male **Ethnicity:** 100% White or Caucasian **Region:** 36% Rural, 28% Suburban, 36% Urban **Teaching Experience:** 7% 1 to 5 years, 36% 6 to 10 years, 21% 11 to 15 years, 7% 16 to 20 years, 29% More than 20 years	66	58
Rhode Island	October ‘22	Virtual	20	**Gender:** 85% Female, 15% Male **Ethnicity:** 10% Black or African American, 5% Multiracial or Biracial, 85% White or Caucasian **Region:** 10% Rural, 25% Suburban, 30% Urban, 35% Did Not Respond **Teaching Experience:** 20% 1 to 5 years, 15% 6 to 10 years, 10% 11 to 15 years, 25% 16 to 20 years, 30% More than 20 years	115	93
South Dakota	October ‘19	Pierre, SD	26	**Gender:** 81% Female, 19% Male **Ethnicity:** 4% American Indian or Alaska Native, 4% Asian, 92% White or Caucasian **Region:** 65% Rural, 15% Suburban, 15% Urban, 4% Not Applicable **Teaching Experience:** 12% 1 to 5 years, 12% 6 to 10 years, 19% 11 to 15 years, 19% 16 to 20 years, 38% More than 20 years	235	222
Utah	July ‘17	Park City, UT	18	**Gender:** 74% Female, 26% Male **Ethnicity:** 4% Native American, 91% White or Caucasian, 4% Other **Teaching Experience:** 100% General Education, 4% Special Education, 4% Other	55	51
	December ‘17	Salt Lake City, UT	36	**Gender:** 84% Female, 16% Male **Ethnicity:** 3% American Indian/Alaska Native and White, 94% White or Caucasian, 3% Other **Teaching Experience:** 87% General Education, 10% General Education and Other, 3% General Education and ESOL	64	62
	October ‘19	Provo, UT	16	**Gender:** 25% Female, 75% Did Not Respond **Ethnicity**: 25% White or Caucasian, 75% Did Not Respond **Region:** 25% Suburban, 75% Did Not Respond	91	44

State/Item Bank	Date	Location	Number of Committee Members	CAC Committee Member Demographic Summary	Number of Items Reviewed	Number of Items Approved by Committees
				Teaching Experience: 6% 6 to 10 years, 6% 16 to 20 years, 13% More than 20 years, 75% Did Not Respond		
	July ‘20[a]	Virtual	17	**State:** 6% Connecticut, 12% Hawaii, 18% Idaho, 12% Montana, 12% Oregon, 12% Rhode Island, 6% Vermont, 6% West Virginia, 12% Wyoming, 6% Did Not Respond **Gender:** 82% Female, 18% Male **Ethnicity**: 47% White or Caucasian, 6% Other, 47% Did Not Respond **Region:** 6% Rural, 12% Suburban, 6% Urban, 76% Did Not Respond **Teaching Experience:** 12% 6 to 10 years, 6% 11 to 15, 18% 16 to 20 years, 18% More than 20 years, 47% Did Not Respond	44	44
	July ‘20	Virtual	16	**Gender:** 31% Female, 6% Male, 63% Did Not Respond **Ethnicity**: 6% Asian or Pacific Islander, 19% Hispanic or Latino, 13% White or Caucasian, 63% Did Not Respond **Region:** 19% Urban, 6% Suburban, 75% Did Not Respond **Teaching Experience:** 6% 6 to 10 years, 13% 11 to 15 years, 13% 16 to 20 years, 6% More than 20 years, 63% Did Not Respond	82	76
	December ‘20	Virtual	6	**Gender:** 50% Female, 50% Did Not Respond **Ethnicity**: 17% Hispanic or Latino, 33% White or Caucasian, 50% Did Not Respond **Region:** 17% Suburban, 83% Did Not Respond **Teaching Experience:** 17% 1 to 5 years, 33% 16 to 20 years, 50% Did Not Respond	14	12
	July ‘21[a]	Virtual	68	**State:** 22% Connecticut, 4% Hawaii, 9% Idaho, 4% Montana, 4% Oregon, 4% Rhode Island, 13% South Dakota, 4% Utah, 3% Vermont, 12% West Virginia, 13% Wyoming, 4% Did Not Respond **Teaching Experience:** 9% 1 to 5 years, 18% 6 to 10 years, 29% 11 to 15 years, 19% 16 to 20 years, 25% More than 20 years	55	53
	August ‘21	Virtual	14	**Gender:** 86% Female, 14% Male **Ethnicity:** 7% Asian or Pacific Islander, 21% Hispanic or Latino, 71% White or Caucasian	62	62

State/Item Bank	Date	Location	Number of Committee Members	CAC Committee Member Demographic Summary	Number of Items Reviewed	Number of Items Approved by Committees
				Region: 14% Rural, 36% Suburban, 43% Urban, 7% Did Not Respond **Teaching Experience:** 7% 1 to 5 years, 21% 6 to 10 years, 21% 11 to 15 years, 29% 16 to 20 years, 21% More than 20 years		
	September '22	Salt Lake City, UT	28	**Gender:** 81% Female, 19% Male **Ethnicity:** 2% American Indian or Alaska Native, 1% Asian, 2% Hispanic or Latino, 1% Mediterranean, 80% White or Caucasian **Region:** 26% Rural, 52% Suburban, 16% Urban, 7% Not Applicable **Teaching Experience:** 29% 1 to 5 years, 39% 6 to 10 years, 3% 11 to 15 years, 7% 16 to 20 years, 7% More than 20 years, 16% Not Applicable	111	99
West Virginia	January '17	Charleston, WV	28[d]	Not Collected	39	[b]
	October '18	Charleston, WV	10	**Gender:** 89% Female, 11% Male **Ethnicity:** 11% Black or African American, 89% White or Caucasian **Region:** 100% Rural **Teaching Experience:** 100% General Education	191	[b]
	January '19	Charleston, WV	9	**Gender:** 89% Female, 11% Male **Ethnicity:** 11% Black or African American, 89% White or Caucasian **Region:** 100% Rural	71	67
	July '19	Charleston, WV	12	**Gender:** 87% Female, 13% Male **Ethnicity:** 4% Asian, 4% Black or African American, 87% White or Caucasian, 4% Not Applicable **Region:** 70% Rural, 30% Urban, 4% Not Applicable **Teaching Experience:** 72% General Education, 4% Special Education, 13% Other, 13% Not Applicable	50	[b]
	July '20[a]	Virtual	8	**State:** 13% Connecticut, 38% Idaho, 13% Oregon, 13% Wyoming, 25% Did Not Respond **Gender:** 100% Female **Ethnicity**: 38% White or Caucasian, 63% Did Not Respond **Region:** 13% Suburban, 13% Rural, 75% Did Not Respond	102	102

State/Item Bank	Date	Location	Number of Committee Members	CAC Committee Member Demographic Summary	Number of Items Reviewed	Number of Items Approved by Committees
				Teaching Experience: 13% 6 to 10 years, 25% More than 20 years, 63% Did Not Respond		
	July ‘21[a]	Virtual	68	**State:** 22% Connecticut, 4% Hawaii, 9% Idaho, 4% Montana, 4% Oregon, 4% Rhode Island, 13% South Dakota, 4% Utah, 3% Vermont, 12% West Virginia, 13% Wyoming, 4% Did Not Respond **Teaching Experience:** 9% 1 to 5 years, 18% 6 to 10 years, 29% 11 to 15 years, 19% 16 to 20 years, 25% More than 20 years	16	16
	June/ August ‘22[a]	Virtual	8	**State:** 25% Connecticut, 13% Hawaii, 25% Idaho, 13% Rhode Island, 25%Wyoming **Gender:** 88% Female, 12% Male **Ethnicity:** 12% Asian or Pacific Islander, 75% White or Caucasian, 12% Did Not Respond **Region:** 25% Suburban, 12% Urban, 63% Did Not Respond **Teaching Experience:** 12% 6 to 10 years, 38% 11 to 15 years, 25% 16 to 20 years, 12% More than 20 years, 12% Did Not Respond	13	13
Wyoming	December ‘17	Cheyenne, WY	18	Not Collected	32	30
	October ‘18	Cheyenne, WY	19	**Gender:** 79% Female, 21% Male **Teaching Experience:** 5% 3 to 5 years, 21% 6 to 10 years, 42% 11 to 20 years, 32% 21 or more years	39	36
	November ‘19	Cheyenne, WY	22	**Gender:** 91% Female, 9% Male **Teaching Experience:** 9% 3 to 5 years, 23% 6 to 10 years, 18% 11 to 20 years, 50% 21 or more years	44	43
	July ‘20[a]	Virtual	13	**State:** 8% Connecticut, 15% Hawaii, 8% Montana, 15% Oregon, 8% Rhode Island, 23% West Virginia, 23% Did Not Respond **Gender:** 77% Female, 23% Male **Ethnicity**: 8% Asian or Pacific Islander, 23% White or Caucasian, 8% Other, 61% Did Not Respond **Region:** 8% Suburban, 15% Urban, 77% Did Not Respond **Teaching Experience:** 15% 6 to 10 years, 23% 11 to 15 years, 15% 16 to 20 years, 46% Did Not Respond	37	37
	August ‘20	Virtual	14	**Gender:** 29% Female, 7% Male, 64% Did Not Respond **Ethnicity**: 36% White or Caucasian, 64% Did Not Respond **Region:** 22% Rural, 78% Did Not Respond	37	36

State/Item Bank	Date	Location	Number of Committee Members	CAC Committee Member Demographic Summary	Number of Items Reviewed	Number of Items Approved by Committees
				Teaching Experience: 7% 11 to 15 years, 7% 16 to 20 years, 22% More than 20 years, 64% Did Not Respond		
	June/ July '21	Virtual	14	**Gender:** 43% Female, 7% Male, 50% Did Not Respond **Ethnicity:** 50% White or Caucasian, 50% Did Not Respond **Region:** 14% Rural,7% Suburban, 7% Urban, 71% Did Not Respond **Teaching Experience:** 14% 11 to 15 years, 36% More than 20 years, 50% Did Not Respond	39	39
	July '21[a]	Virtual	68	**State:** 22% Connecticut, 4% Hawaii, 9% Idaho, 4% Montana, 4% Oregon, 4% Rhode Island, 13% South Dakota, 4% Utah, 3% Vermont, 12% West Virginia, 13% Wyoming, 4% Did Not Respond **Teaching Experience:** 9% 1 to 5 years, 18% 6 to 10 years, 29% 11 to 15 years, 19% 16 to 20 years, 25% More than 20 years	39	38
	June/ August '22[a]	Virtual	9	**State:** 22% Connecticut, 22% Idaho, 33% Utah, 22% West Virginia **Gender:** 78% Female, 22% Male **Ethnicity:** 11% Hispanic or Latino, 78% White or Caucasian, 11% Race/Ethnicity Not Listed **Region:** 33% Rural, 11% Suburban, 11% Urban, 44% Did Not Respond **Teaching Experience:** 44% 6 to 10 years, 11% 11 to 15 years, 44% More than 20 years	37	37

[a]Items were reviewed in a combined Content Advisory Committee Meeting that included ICCR and MOU state-owned items. Items reviewed in the combined meetings are displayed by their respective state or bank of ownership.
[b]As of the time of writing this report, the number of science items reviewed and/or approved by Content Advisory Committees is not currently available.
[c]The number of Committee Members is not available at the time of writing this report.
[d]The number of Committee Members includes the total members across ELA, math, and science committees. The specific number of science committee members is currently unavailable.

Appendix 2-G

Fairness Committee Review Training Slides

<u>Fairness Committee Review Training Slides</u>

Agenda

1. Item Review Training
2. Content Rater Demo
3. Batch Review (~1 hr)
4. Item Discussion

Test Security

- Non-disclosure agreements should have already been signed.
- All test materials viewed during this meeting are considered secure.
- Do NOT discuss test material content outside of this meeting.

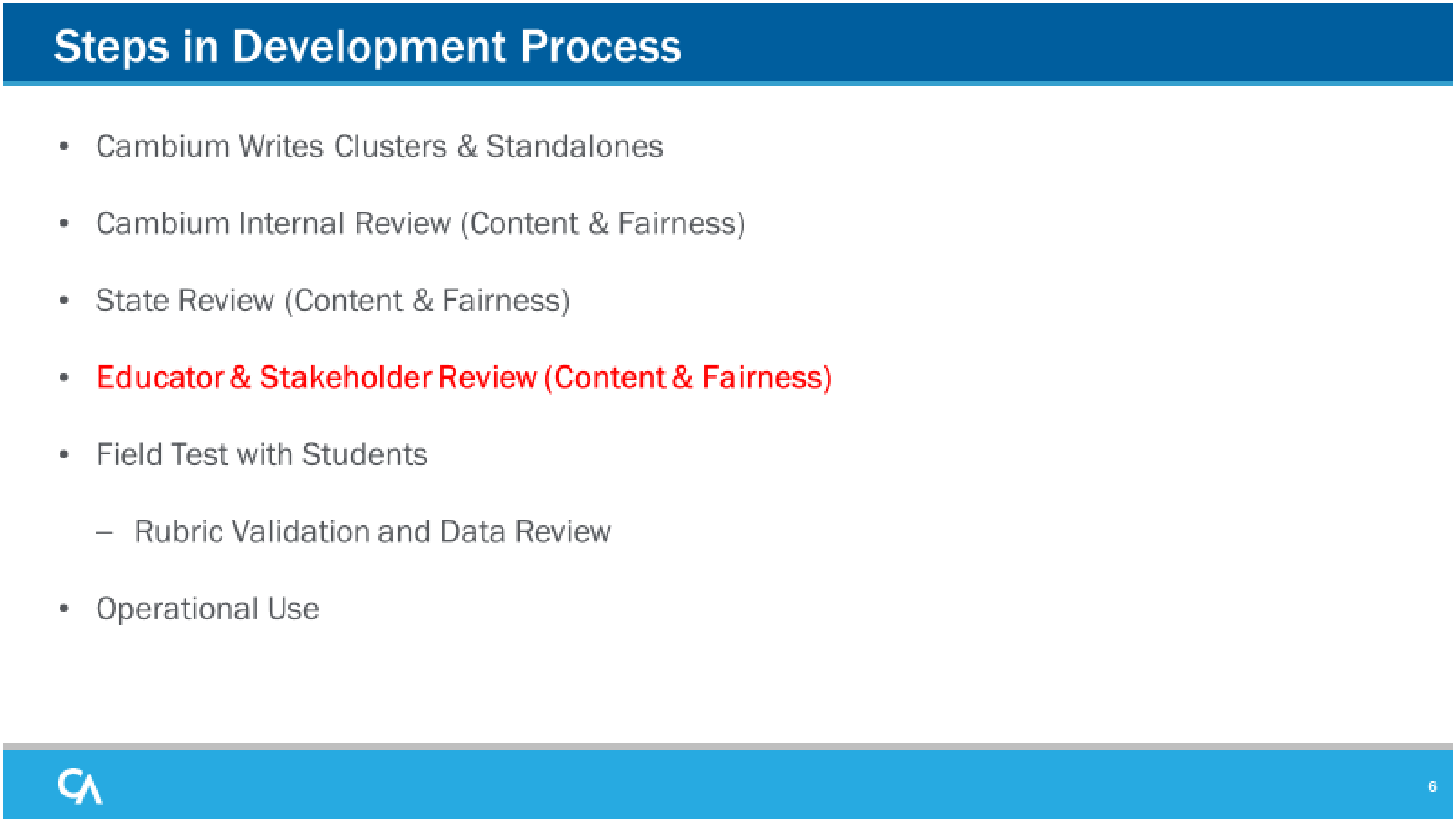
3

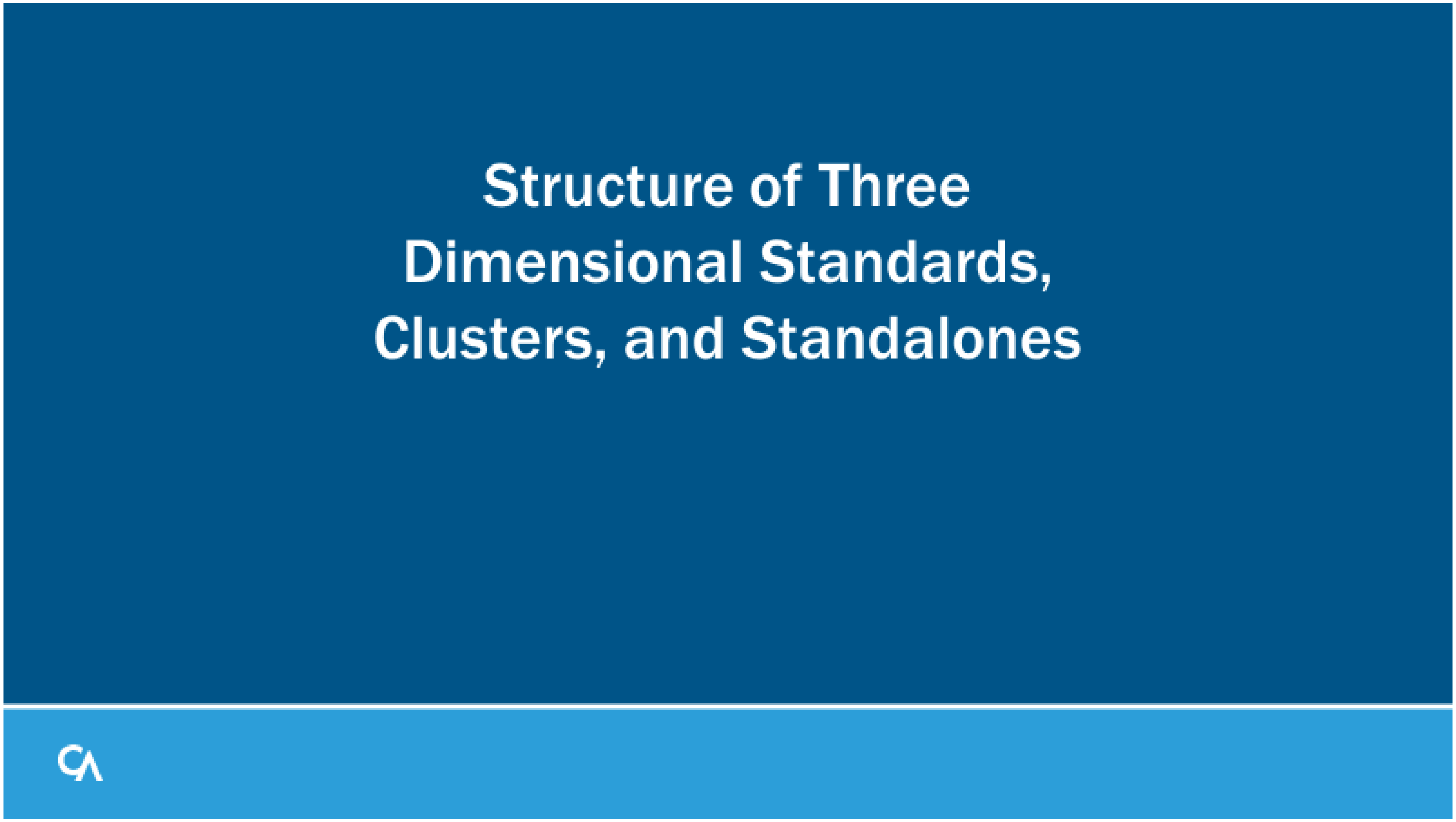
Overview of Item Review Training

- Steps in the Development Process
- Describe Structure of Three-Dimensional Standard, Clusters & Standalones
- Fairness Review Process
- Norms & Participant Guidelines

4


Steps in the Development
Process


Steps in Development Process
Cambium Writes Clusters & Standalones
Cambium Internal Review (Content & Fairness)
State Review (Content & Fairness)
Educator & Stakeholder Review (Content & Fairness)
Field Test with Students
– Rubric Validation and Data Review
Operational Use
6


Structure of Three
Dimensional Standards,
Clusters, and Standalones

Three-Dimensional Standards

- Each 3-D "standard" is a blend of one of several scientific activities that are common to the doing of all sciences (SEP), one or two "big ideas" from a science discipline (DCI), and one of several broad themes that are found across scientific disciplinary boundaries (CCC).

8

Three-Dimensional Standard

Students who demonstrate understanding can:

MS-LS2-4. **Construct an argument supported by empirical evidence that changes to physical or biological components of an ecosystem affect populations.** [Clarification Statement: Emphasis is on recognizing patterns in data and making warranted inferences about changes in populations, and on evaluating empirical evidence supporting arguments about changes to ecosystems.]

The performance expectation above was developed using the following elements from the NRC document *A Framework for K-12 Science Education*:

Science and Engineering Practices

Engaging in Argument from Evidence
Engaging in argument from evidence in 6–8 builds on K–5 experiences and progresses to constructing a convincing argument that supports or refutes claims for either explanations or solutions about the natural and designed world(s).

- Construct an oral and written argument supported by empirical evidence and scientific reasoning to support or refute an explanation or a model for a phenomenon or a solution to a problem.

Connections to Nature of Science

Scientific Knowledge is Based on Empirical Evidence

- Science disciplines share common rules of obtaining and evaluating empirical evidence.

Disciplinary Core Ideas

LS2.C: Ecosystem Dynamics, Functioning, and Resilience

- Ecosystems are dynamic in nature; their characteristics can vary over time. Disruptions to any physical or biological component of an ecosystem can lead to shifts in all its populations.

Crosscutting Concepts

Stability and Change

- Small changes in one part of a system might cause large changes in another part.

Connections to other DCIs in this grade-band:
MS.LS4.C ; **MS.LS4.D** ; **MS.ESS2.A** ; **MS.ESS3.A** ; **MS.ESS3.C**

Articulation of DCIs across grade-bands:
3.LS2.C ; **3.LS4.D** ; **HS.LS2.C** ; **HS.LS4.C** ; **HS.LS4.D** ; **HS.ESS2.E** ; **HS.ESS3.B** ; **HS.ESS3.C**

Common Core State Standards Connections:
ELA/Literacy -

RST.6-8.1	Cite specific textual evidence to support analysis of science and technical texts. (MS-LS2-4)
RI.8.8	Trace and evaluate the argument and specific claims in a text, assessing whether the reasoning is sound and the evidence is relevant and sufficient to support the claims. (MS-LS2-4)
WHST.6-8.1	Write arguments to support claims with clear reasons and relevant evidence. (MS-LS2-4)
WHST.6-8.9	Draw evidence from literary or informational texts to support analysis, reflection, and research. (MS-LS2-4)

9

Three Dimensions

Scientific and Engineering Practices	Disciplinary Core Ideas	Cross Cutting Concepts
Asking Questions and Defining Problems	Physical Science	Patterns
Developing and Using Models	Life Science	Cause and Effect
Planning and Carrying Out Investigations	Earth and Space Science	Scale, Proportion, and Quantity
Analyzing and Interpreting Data	Engineering, Technology, and the Application of Science	Systems and System Models
Using Mathematics and Computational Thinking		Energy and Matter
Constructing Explanations and Designing Solutions		Structure and Function
Engaging in Argument from Evidence		Stability and Change
Obtaining, Evaluating, and Communicating Information		

10

Structure of CAI Clusters

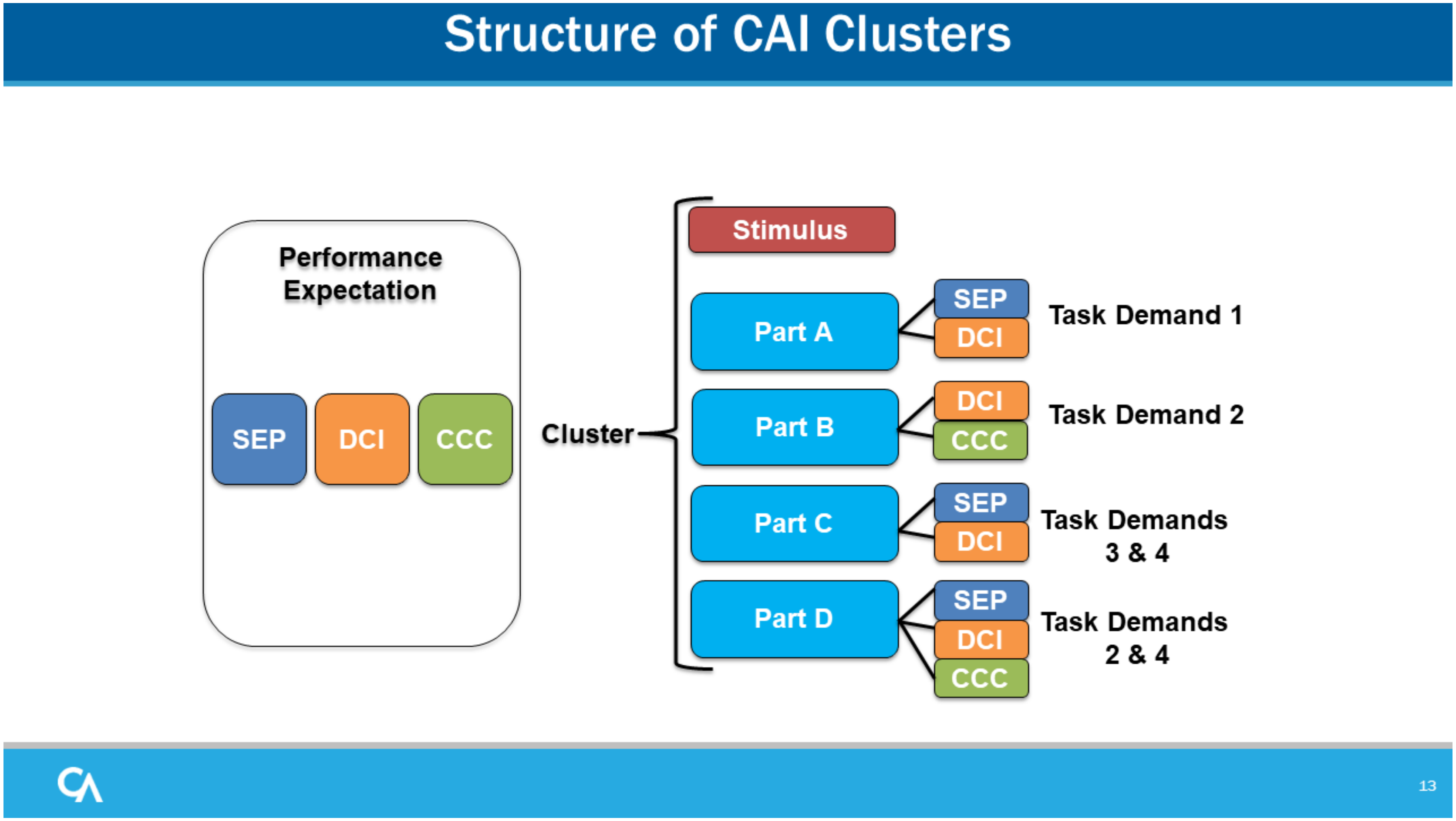
- Each cluster begins with a **phenomenon,** which is the observation about the natural world which anchors the entire cluster. The interactions within the cluster all address the phenomenon.
- Each cluster engages the student in a grade-appropriate, meaningful **scientific activity** aligned to a specific standard.
- A **cluster task statement** comes at the end of the stimulus and provides an overview of the point of the cluster.
- Each interaction in the cluster **aligns** to at least two of the three dimensions (SEP, DCI, CCC) and, if possible, all three.

11

Structure of Cambium Clusters

Task Demands
1. Based on the provided data, identify, describe or illustrate a claim regarding the relationship between a characteristic animal behavior and/or specialized plant structure and the probability of successful reproduction in the species.
2. Identify, summarize, or organize given data or other information to support or refute a claim regarding the relationship between a characteristic animal behavior and/or specialized plant structure and the probability of successful reproduction in the species.
3. Sort inferences about the relationship of behaviors or structures to breeding success into those that are supported by the data, contradicted by the data, or neither, or some similar classification.
4. Select supporting evidence from competing sources based on the reliability of statistical relationships, how representative the sample is, or study design.
5. Construct an argument using scientific reasoning drawing on credible evidence to explain the relationships of animal behaviors or plant structures to reproductive success. (Hand scored CR)
6. Identify additional evidence that would help clarify, support, or contradict a hypothesized relationship or causal argument.
7. Identify or describe alternate explanations and the data needed to distinguish among them.

12


Structure of CAI Clusters
Performance Expectation
SEP
DCI
CCC
Cluster
Stimulus
Part A
SEP
DCI
Task Demand 1
Part B
DCI
CCC
Task Demand 2
Part C
SEP
DCI
Task Demands 3 & 4
Part D
SEP
DCI
CCC
Task Demands 2 & 4
13

Willow populations in Yellowstone National Park have increased since wolves were reintroduced to the park in 1995.

Willows are small trees that grow best in marshlike environments. After studying the Yellowstone food web shown in Diagram 1 and the population data for the park shown in Table 1, students arrive at two different hypotheses.

Diagram 1. Yellowstone Food Web

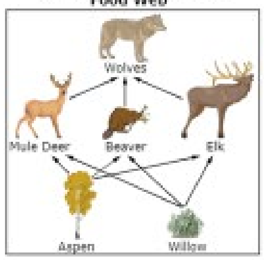


Table 1. Yellowstone Population Data

	Wolves	Elk	Beaver	Mule Deer
1995	31	16,791	10	2,014
2004	171	8,335	120	2,014

Note: These data are approximate.

Hypothesis 1:
When wolves were reintroduced to Yellowstone, the wolves preyed upon the elk, which allowed the beavers to eat more willow. This led to more beavers and beaver dams. Beaver dams create marsh environments that willows do well in, allowing the willow's population to increase.

Part A

Click on each box and select a word/phrase that completes the table with the Yellowstone population data from 1995 and 2004 and the hypothesis those data support.

Table 2. Summary of Yellowstone Population Data and Supported Hypotheses

Data	Hypothesis Supported
Elk population ▾	▾
Beaver population ▾	▾
Mule deer population ▾	▾

Part B

Which hypothesis is best supported by the evidence?

Ⓐ All of the evidence is consistent with Hypothesis 1.

Ⓑ All of the evidence is consistent with Hypothesis 2.

Ⓒ Most of the evidence is consistent with Hypothesis 1.

Ⓓ Most of the evidence is consistent with Hypothesis 2.

Ⓔ The evidence does not favor either hypothesis.

Part C

Aspen trees are shown in Diagram 1. Moose and bison are two plant-eating animal species that are not shown in Diagram 1 but are also part of the Yellowstone food web.

Based on Hypothesis 2, click on each box to select a word/phrase to make a prediction about what would happen to the moose, bison, and aspen tree populations after the reintroduction of wolves.

Table 3. Population Predictions

Species	Population after Wolf Reintroduction	Reason for Impact on Population
Moose	▾	▾
Bison	▾	▾
Aspen tree	▾	▾

14

Scoring Assertions

- Within each item are a series of explicit assertions that can be made about the knowledge and skills that a student has demonstrated based on specific features of the student's responses
- Scoring assertions tell us:
  - What the student did
  - What inference can be made from that action
- Elementary clusters have 5 or more assertions. Middle and High school clusters have 7 or more.

15

Scoring Assertions

Your response earned **0** points of a possible **15**

Score Rationale	
The student identified that the elk population decreased between 1995 and 2004 giving evidence of understanding of how to interpret the data presented in Table 1.	✗
The student identified that the elk population decreased between 1995 and 2004 and that data supports both hypotheses giving evidence of understanding of how to use evidence to support an argument.	✗
The student identified that the beaver population increased between 1995 and 2004 giving evidence of understanding of how to interpret the data presented in Table 1.	✗
The student identified that the beaver population increased between 1995 and 2004 and that data supports Hypothesis 1 giving evidence of understanding of how to use evidence to support an argument.	✗
The student identified that the mule deer population had no change between 1995 and 2004 giving evidence of understanding of how to interpret the data presented in Table 1.	✗
The student identified that the mule deer population had no change between 1995 and 2004 and that data supports neither hypothesis giving evidence of understanding of how to use evidence to support an argument.	✗
The student identified that most of the evidence is consistent with Hypothesis 1 or the student's conclusion matches their correct inferences, given their characterization of the data, providing some evidence of his or her ability to summarize information and draw a conclusion.	✗
The student predicted that the moose population would decrease giving evidence of the ability to extrapolate information from a given food web.	✗

16

Standalone Items

- Standalone items are single-interaction items that are closer to traditional assessment items.
- Each standalone item is at least 2-dimensionally aligned (SEP + DCI), if not 3-dimensional.
- Standalone items (all grades) contain between 1 and 4 assertions.

17


Fairness Review Process

Assessment Fairness

Purpose:

- Test materials need to be free from unnecessary barriers to the success of diverse groups of students.

Why?

1. Valid measurement of student achievement.
2. Positive student experience.

19

Assessment Fairness

Assessment content is...

1. free of bias and stereotypes.
2. sensitive to student and community beliefs and experiences.
3. accessible to all students, to the greatest extent possible.

20

Assessment Fairness

Assessment content is free of bias and stereotypes.

- Assessment content does not express bias or present stereotypes of people of different groups.
- Assessment content demonstrates respect for people of different groups.
- Assessment content is inclusive and reflects the diversity of the community.

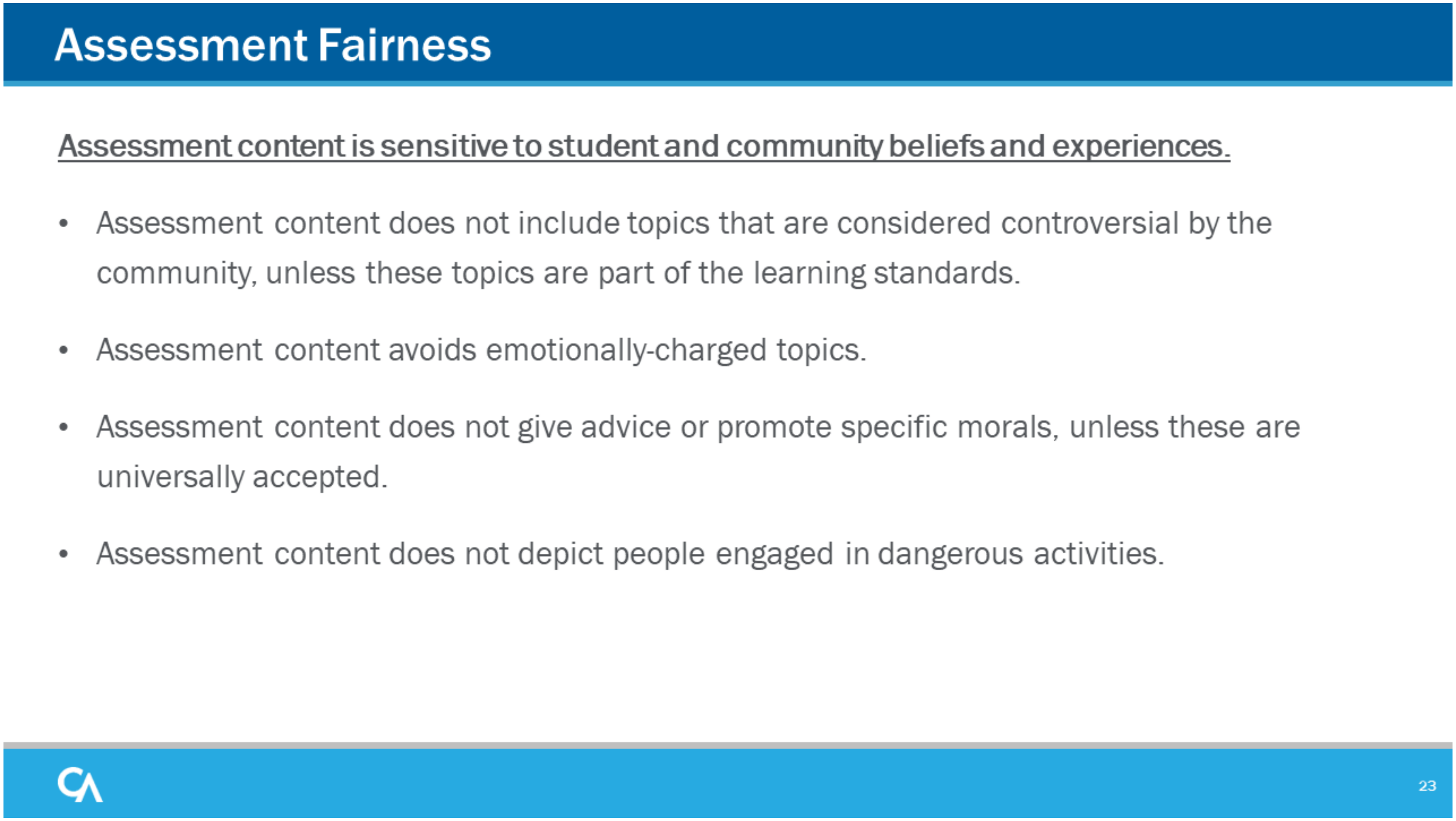
21

Assessment Fairness

Assessment content is free of bias and stereotypes.

- Gender
- Race, Ethnicity, Culture
- Religion
- Age
- Disability
- Socioeconomic

22


Assessment Fairness
Assessment content is sensitive to student and community beliefs and experiences.
• Assessment content does not include topics that are considered controversial by the community, unless these topics are part of the learning standards.
• Assessment content avoids emotionally-charged topics.
• Assessment content does not give advice or promote specific morals, unless these are universally accepted.
• Assessment content does not depict people engaged in dangerous activities.
23

Assessment Fairness

Assessment content is sensitive to student and community beliefs and experiences.

- Does the material require a student to take a position that challenges authority?
- Does the material present sensitive or highly controversial subjects, such as death, war, abortion, euthanasia, or natural disasters, except where they are needed to assess learning standards?
- Does the material trivialize significant or tragic human experiences?
- Does the material require the parent, teacher, or examinee to support a position that is contrary to their religious beliefs?
- Does the material contain advice pertaining to health and well-being about which there is not universal agreement?
- Does the material portray people engaged in dangerous activities, unless required to assess the learning standard?

24

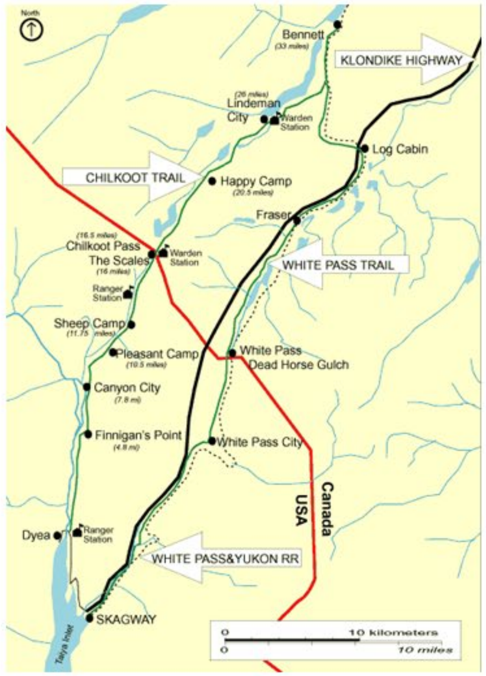
Assessment Fairness

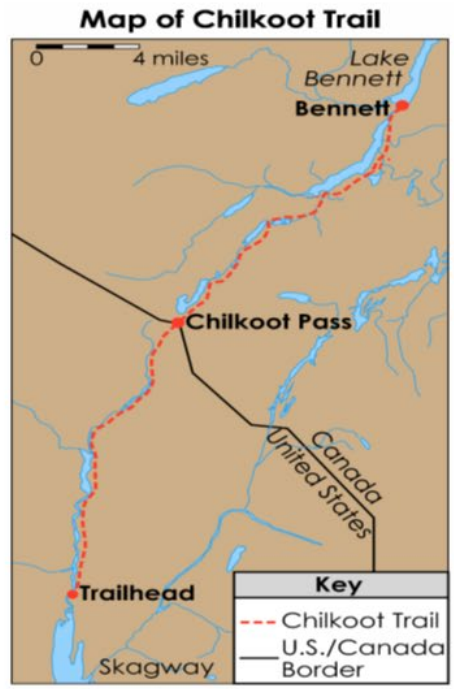
Assessment content is accessible to all students to the greatest extent possible.

- Language used in the assessment is understandable and inclusive.
  - a. Does the material use exclusive or gender-specific terms rather than inclusive terms?
  - b. Does the material use words or phrases that may not be similarly understood by students of different groups?
- Assessment content does not rely on vocabulary or background knowledge that would be significantly more or less familiar to a student because of their demographic group. (Differential familiarity)
  - a. Does the material contain phrases, concepts, and beliefs that are irrelevant to testing domain and are likely to be more familiar to specific groups than others?
  - b. Does the material require knowledge of individuals, events, or groups that is not familiar to all groups of students?
- Accommodations and supports are available to students when needed.
  - a. Is the content incompatible with the accommodations or supports provided to students?
    - » Graphic complexity & color

25

Assessment Fairness





26

Fairness Review Process

1. Enter Content Rater system and begin batch.
2. Review fairness of each item.
3. Document comments and answers to review questions in content rater for discussion

27

Fairness Review Process

- Cambium facilitator will lead discussion of items.
  - Facilitator documents group decision/feedback from committee directly into ITS.
    - » Specific fairness concern / issue
    - » Suggested revisions
  - Cambium facilitator also documents review outcome recommendation in ITS.
    - » AAA-Accept as Appears
    - » AAR-Accept as Revised
    - » Recommend Reject
- Later, feedback will be discussed with individual state departments of education.

28

Fairness Review Procedures

1. Accept or Reject from a fairness perspective
   - Use Fairness Criteria Guidelines during review
   - Record notes in content rater
2. Facilitator will project items and record feedback by pending committee comments directly into ITS
   - Specific fairness concern / issue
   - Suggested revisions

29

Norms and Participant Guidelines

Participant Guidelines

- Participants are required to be on camera throughout all sessions.
  - This includes during introductions and group discussions. The session facilitator will signal when participants should turn cameras on and off.
- Keep phones turned off & stowed while completing review.
- Utilize Go-To meeting chat for any questions or concerns during the individual review time.

31

Norms

- Listen actively and attentively.
- Ask for clarification if you are confused.
- Avoid put-downs (even humorous ones).
- Build on one another's comments; work toward shared understanding.
- Avoid off topic & side conversations.

32

Item Security

- It is imperative that participants maintain security of the items, test content, and information included in these sessions.
  - All participants have been required to sign NDAs before these sessions and have the responsibility to adhere to the requirements of those agreements for these sessions.
  - The NDAs explain that the items, test content, test data, technology details and other proprietary information included in these sessions are confidential information that must be kept secure.
  - This confidential information should not be disclosed to or discussed with anyone outside of these sessions. Participants are required to notify CAI if any security breach occurs-i.e., any unauthorized use or disclosure of the confidential information.
- Be aware of other people who may hear the discussion or see your computer screen. Item content must remain secure.

33


Questions?

Appendix 2-H

Fairness Committee Participant Details

Fairness Committee Participant Details

Table H-1. Fairness Committee Participants, Science

State/Item Bank	Date	Location	Number of Committee Members	Fairness Committee Member Demographic Summary	Number of Items Reviewed	Number of Items Rejected by Committees
Connecticut	February ‘17	Cromwell, CT	6	**Gender:** 83% Female, 17% Male	45	1
	December ‘17	New Britain, CT	9	**Gender:** 78% Female, 22% Male	75	[a]
	December ‘17	Cromwell, CT	10	**Gender:** 70% Female, 30% Male	41	[a]
	February ‘18	New Britain, CT	3	**Gender:** 67% Female, 33% Male	42	1
	November ‘18	New Britain, CT	11	**Gender:** 91% Female, 9% Male	319	38
	December ‘18	New Britain, CT	10	**Gender:** 80% Female, 20% Male	56	1
	January ‘19	New Britain, CT	9	**Gender:** 78% Female, 22% Male	65	1
	September ‘19	Cromwell, CT	9	**Gender:** 89% Female, 11% Male	48	0
	November ‘19	Cromwell, CT	10	**Gender:** 80% Female, 20% Male **Ethnicity:** 100% White or Caucasian **Region:** 10% Rural, 70% Suburban, 20% Urban **Teaching Experience:** 10% 6 to 10 years, 20% 11 to 15 years, 10% 16 to 20 years, 60% More than 20 years	52	1
	July ‘20[b]	Virtual	8	**Gender:** 88% Female, 13% Male **Ethnicity:** 13% Hispanic or Latino, 75% White or Caucasian, 13% Prefer Not to Say	43	0
	July ‘21[b]	Virtual	6	**State:** 22% Connecticut, 4% Hawaii, 9% Idaho, 4% Montana, 4% Oregon, 4% Rhode Island, 13% South Dakota, 4% Utah, 3% Vermont, 12% West Virginia, 13% Wyoming, 4% Did Not Respond **Teaching Experience:** 9% 1 to 5 years, 18% 6 to 10 years, 29% 11 to 15 years, 19% 16 to 20 years, 25% More than 20 years	20	0
	September ‘21	Virtual	7	**Gender:** 43% Female, 57% Male **Ethnicity:** 100% White or Caucasian	111	23

State/Item Bank	Date	Location	Number of Committee Members	Fairness Committee Member Demographic Summary	Number of Items Reviewed	Number of Items Rejected by Committees
				Region: 29% Suburban, 29% Urban, 43% Did Not Respond **Teaching Experience:** 14% 6 to 10 years, 29% 11 to 15 years, 14% 16 to 20 years, 43% More than 20 years		
	June/ August ‘22[b]	Virtual	3	**State:** 33% Connecticut, 33% Oregon, 33% Wyoming **Gender:** 66% Female, 33% Male **Ethnicity:** 66% White or Caucasian, 33% Did Not Respond **Region:** 66% Suburban, 33% Did Not Respond **Teaching Experience:** 66% More than 20 years, 33% Did Not Respond	65	2
	August ‘22	Virtual	19	**Gender:** 79% Female, 21% Male **Ethnicity:** 5% Hispanic or Latino, 90% White or Caucasian, 5% Multiracial or Biracial **Region:** 63% Suburban, 5% Urban, 32% Not Applicable **Teaching Experience:** 11% 6 to 10 years, 11% 11 to 15 years, 21% 16 to 20 years, 57% More than 20 years	154	27
Hawaii	July ‘17	Honolulu, HI	22	**Gender:** 64% Female, 36% Male **Ethnicity:** 5% Black or African American, 5% Chinese and White, 9% Filipino, 14% Hawaiian, 9% Hispanic or Latino, 14% Japanese, 41% White or Caucasian, 5% Did Not Respond **Teaching Experience:** 64% General Education, 5% General Education with SPED Certification, 5% SPED Teacher, 23% Other, 5% Did Not Respond	25	2
	September ‘17	Honolulu, HI	20	**Gender:** 75% Female, 25% Male **Ethnicity:** 5% Black or African American, 10% Filipino, 10% Hispanic or Latino, 15% Japanese, 50% White or Caucasian, 10% Did Not Respond **Teaching Experience:** 65% General Education, 15% General Education with SPED Certification, 20% Other	65	13
	October ‘18	Honolulu, HI	29	**Gender:** 79% Female, 21% Male **Ethnicity:** 7% Asian, 3% Hawaiian, 7% Asian Pacific Islander, 7% Chinese, 3% Filipino, 10% Hispanic or Latino, 10% Japanese, 28% White or Caucasian, 14% Multi-Racial/Ethnic, 10% Not Applicable	85	6
	February/ March ‘19	Honolulu, HI	21	**Gender:** 80% Female, 20% Male **Ethnicity:** 50% Asian, 35% White or Caucasian, 15% Two or More	44	0

State/Item Bank	Date	Location	Number of Committee Members	Fairness Committee Member Demographic Summary	Number of Items Reviewed	Number of Items Rejected by Committees
				Teaching Experience: 65% General Education, 5% General Education with SPED Certification, 5% SPED Teacher, 25% Other		
	June/ July '20	Virtual	17	**Gender:** 18% Female, 12% Male, 70% Did Not Respond **Ethnicity**: 18% White or Caucasian, 6% Asian or Pacific Islander, 6% Multiracial or Biracial, 70% Did Not Respond **Region:** 12% Rural, 12% Suburban, 76% Did Not Respond **Teaching Experience:** 6% 6 to 10 years, 12% 11 to 15 years, 12% More than 20 years, 70% Did Not Respond	344	324
	July '20[b]	Virtual	4	**State:** 25% Connecticut, 50% Rhode Island, 25%Utah **Gender:** 100% Female **Ethnicity**: 25% White or Caucasian, 25% Hispanic or Latino, 50% Did Not Respond **Region:** 25% Urban, 75% Did Not Respond **Teaching Experience:** 25% 6 to 10 years, 25% 16 to 20 years, 50% Did Not Respond	55	8
	June/ August '22[b]	Virtual	6	**State:** 33% Oregon, 17% Rhode Island, 33% Utah, 17% West Virginia **Gender:** 67% Female, 33% Male **Ethnicity:** 67% White or Caucasian, 33% Multiracial or Biracial **Region:** 33% Rural, 33% Urban, 33% Did Not Respond **Teaching Experience:** 33% None, 33% 11 to 15 years, 33% More than 20 years	46	0
	July '22	Honolulu, HI	9	**Gender:** 67% Female, 22% Male, 11% Non-Binary **Ethnicity:** 44% Asian/Pacific Islander, 11% Hispanic or Latino, 11% Multiracial or Biracial, 22% White or Caucasian, 11% Did Not Respond **Region:** 11% Hilo, 11% Maui, 78% Oahu **Teaching Experience:** 11% 1 to 5 years, 11% 6 to 10 years, 11% 11 to 15 years, 22% 16 to 20 years, 44% More than 20 years	45	0
	July '22	Honolulu, HI	9	**Gender:** 67% Female, 22% Male, 11% Non-Binary **Ethnicity:** 44% Asian/Pacific Islander, 11% Hispanic or Latino, 11% Multiracial or Biracial, 22% White or Caucasian, 11% Did Not Respond **Region:** 11% Hilo, 11% Maui, 78% Oahu	306	0

State/Item Bank	Date	Location	Number of Committee Members	Fairness Committee Member Demographic Summary	Number of Items Reviewed	Number of Items Rejected by Committees
				Teaching Experience: 11% 1 to 5 years, 11% 6 to 10 years, 11% 11 to 15 years, 22% 16 to 20 years, 44% More than 20 years		
ICCR	March '18	Virtual	13	**State:** 46% Connecticut, 8% Indiana, 15% Utah, 23% West Virginia, 8% Wyoming **Gender:** 85% Female, 15% Male	152	7
	July '20[b]	Virtual	5	**State:** 20% Connecticut, 40% Rhode Island, 20% Utah, 20% Vermont **Gender:** 100% Female **Ethnicity**: 60% White or Caucasian, 20% Hispanic or Latino, 20% Did Not Respond **Region:** 40% Rural, 20% Suburban, 20% Urban, 20% Did Not Respond **Teaching Experience:** 20% 6 to 10 years, 20% 11 to 15 years, 20% 16 to 20 years, 20% More than 20 years, 20% Did Not Respond	57	0
	July '21[b]	Virtual	15	**State:** 22% Connecticut, 4% Hawaii, 9% Idaho, 4% Montana, 4% Oregon, 4% Rhode Island, 13% South Dakota, 4% Utah, 3% Vermont, 12% West Virginia, 13% Wyoming, 4% Did Not Respond Teaching Experience: 9% 1 to 5 years, 18% 6 to 10 years, 29% 11 to 15 years, 19% 16 to 20 years, 25% More than 20 years	157	1
	June/ August '22[b]	Virtual	7	**State:** 71% Connecticut, 14% North Dakota, 14% Utah **Gender:** 100% Female **Ethnicity:** 71% White or Caucasian, 29% Did Not Respond **Region:** 29% Suburban, 71% Did Not Respond **Teaching Experience:** 14% 11 to 15 years, 29% 16 to 20 years, 29% More than 20 years, 29% Did Not Respond	121	3
Idaho	December '18	Boise, ID	15	Not Collected	111	1
	December '21	Boise, ID	21	**Gender:** 81% Female, 19% Male **Ethnicity**: 95% White or Caucasian, 5% Hispanic or Latino **Region:** 33% Rural, 19% Suburban, 5% Urban, 43% Did Not Respond **Teaching Experience:** 19% None, 5% Less than 1 year, 5% 1 to 5 years, 19% 6 to 10 years, 5% 11 to 15 years, 14% 16 to 20 years, 33% More than 20 years	179	0
	June/ August '22[b]	Virtual	4	**State:** 50% Connecticut, 25% Oregon, 25% Wyoming	12	0

State/Item Bank	Date	Location	Number of Committee Members	Fairness Committee Member Demographic Summary	Number of Items Reviewed	Number of Items Rejected by Committees
				Gender: 100% Female **Ethnicity:** 75% White or Caucasian, 25% Did Not Respond **Region:** 25% Suburban, 75% Did Not Respond **Teaching Experience:** 25% 11 to 15 years, 25% 16 to 20 years, 25% More than 20 years, 25% Did Not Respond		
Montana	January ‘20	Helena, MT	15	Not Collected	149	[a]
	July ‘21[b]	Virtual	3	**State:** 22% Connecticut, 4% Hawaii, 9% Idaho, 4% Montana, 4% Oregon, 4% Rhode Island, 13% South Dakota, 4% Utah, 3% Vermont, 12% West Virginia, 13% Wyoming, 4% Did Not Respond **Teaching Experience:** 9% 1 to 5 years, 18% 6 to 10 years, 29% 11 to 15 years, 19% 16 to 20 years, 25% More than 20 years	41	0
	June/ August ‘22[b]	Virtual	4	**State:** 50% Connecticut, 25% Oregon, 25% Wyoming **Gender:** 100% Female **Ethnicity:** 75% White or Caucasian, 25% Did Not Respond **Region:** 25% Suburban, 75% Did Not Respond **Teaching Experience:** 25% 11 to 15 years, 25% 16 to 20 years, 25% More than 20 years, 25% Did Not Respond	13	0
Multi-State Science Assessment (Rhode Island and Vermont)	January ‘18	Providence, RI	21	**State:** 100% Rhode Island **Teaching Experience:** 67% General Education, 14% Bilingual Education, 5% Special Education, 5% Science Coordinator, 10% Other	73	14
	March ‘18	Providence, RI	11	**State:** 55% Rhode Island, 45% Vermont	100	24
	January ‘19	Concord, NH	14	**Gender:** 63% Female, 23% Male **Teaching Experience:** 69% General Education, 3% Special Education, 11% Coach, 17% Other	116	18
	November ‘19	Fairlee, VT	17	**State:** 29% Rhode Island, 6% Vermont, 65% Did Not Respond **Gender:** 23% Female, 12% Male, 65% Did Not Respond **Ethnicity**: 35% White or Caucasian, 65% Did Not Respond **Region:** 6% Rural, 17% Suburban, 77% Did Not Respond **Teaching Experience:** 6% 11 to 15 years, 17% 16 to 20 years, 12% More than 20 years, 65% Did Not Respond	66	0
	July ‘20[b]	Virtual	2	**State:** 50% Utah, 50% Vermont **Gender:** 100% Female **Ethnicity**: 50% Hispanic or Latino, 50% White or Caucasian	27	0

State/Item Bank	Date	Location	Number of Committee Members	Fairness Committee Member Demographic Summary	Number of Items Reviewed	Number of Items Rejected by Committees
				Region: 50% Rural, 50% Did Not Respond **Teaching Experience:** 50% 6 to 10 years, 50% More than 20 years		
	July ‘21[b]	Virtual	3	**State:** 22% Connecticut, 4% Hawaii, 9% Idaho, 4% Montana, 4% Oregon, 4% Rhode Island, 13% South Dakota, 4% Utah, 3% Vermont, 12% West Virginia, 13% Wyoming, 4% Did Not Respond **Teaching Experience:** 9% 1 to 5 years, 18% 6 to 10 years, 29% 11 to 15 years, 19% 16 to 20 years, 25% More than 20 years	30	1
	August ‘21	Virtual	3	**State:** 100% Rhode Island **Gender:** 100% Female **Ethnicity:** 100% White or Caucasian **Region:** 33% Suburban, 67% Urban **Teaching Experience:** 33% 6 to 10 years, 67% More than 20 years	93	3
Oregon	September ‘17	Salem, OR	5	**Gender:** 100% Female **Region:** 80% Urban, 20% Suburban **Teaching Experience:** 40% General Education, 20% Bilingual Education, 20% Special Education, 60% Administration, 20% Other	235	114
	August ‘18	Salem, OR	39	**Gender:** 74% Female, 26% Male **Ethnicity:** 3% Asian, 8% Hispanic or Latino, 3% Native American, 82% White or Caucasian, 10% Other **Region:** 56% Urban, 44% Rural **Teaching Experience:** 15% General Education, 72% Bilingual Education, 33% Special Education, 33% Other	257	8
	October ‘18	Salem, OR	8	**Gender:** 100% Female **Ethnicity:** 80% White or Caucasian, 20% Other **Region:** 80% Urban, 20% Rural **Teaching Experience:** 88% Bilingual Education, 50% Special Education, 63% Other	60	12
	December ‘18	Virtual	11	**Gender:** 91% Female, 9% Male **Ethnicity:** 9% Hispanic or Latino, 91% White or Caucasian **Region:** 55% Urban, 45% Rural **Teaching Experience:** 27% General Education, 64% Bilingual Education, 18% Special Education, 9% Administration, 64% Other	62	14

State/Item Bank	Date	Location	Number of Committee Members	Fairness Committee Member Demographic Summary	Number of Items Reviewed	Number of Items Rejected by Committees
	October '19	Salem, OR	9	**Gender:** 78% Female, 22% Male **Ethnicity:** 89% White or Caucasian, 11% Native American **Region:** 44% Urban, 56% Rural **Teaching Experience:** 89% General Education, 67% Bilingual Education, 44% Special Education	246	23
	January '20	Salem, OR	11	**Gender:** 55% Female, 45% Male **Ethnicity:** 100% White or Caucasian **Region:** 45% Urban, 45% Suburban, 9% Rural **Teaching Experience:** 100% General Education, 90% Bilingual Education, 81% Special Education, 81% Other	262	33
	July '20[b]	Virtual	2	**State:** 50% Connecticut, 50% Utah **Gender:** 100% Female **Ethnicity**: 50% Hispanic or Latino, 50% Did Not Respond **Region:** 100% Did Not Respond **Teaching Experience:** 50% 6 to 10 years, 50% Did Not Respond	22	3
	August '20	Virtual	7	**Gender:** 72% Female, 14% Male, 14% Nonbinary **Ethnicity:** 14% Asian, 43% African American, 29% Hispanic or Latino, 14% Native American **Region:** 14% Urban, 72% Suburban, 14% Rural **Teaching Experience:** 57% General Education, 57% Bilingual Education, 29% Special Education, 29% Administration	86	7
	August '21	Virtual	7	**Gender:** 100% Female **Ethnicity:** 100% White or Caucasian **Region:** 14% Urban, 29% Suburban, 57% Rural **Teaching Experience:** 43% General Education, 14% Bilingual Education, 14% Administration, 29% Other	353	13
	July '22	Virtual	9	**Gender:** 56% Female, 33% Male, 11% Nonbinary **Ethnicity:** 11% Hispanic or Latino, 89% White or Caucasian **Region:** 22% Rural, 45% Suburban, 33% Urban **Teaching Experience:** 23% 1 to 5 years, 33% 6 to 10 years, 11% 11 to 15 years, 33% More than 20 years	43	2
Rhode Island	October '22	Virtual	20	**Gender:** 85% Female, 15% Male **Ethnicity:** 10% Black or African American, 5% Multiracial or Biracial, 85% White or Caucasian **Region:** 10% Rural, 25% Suburban, 30% Urban, 35% Did Not Respond	115	22

State/Item Bank	Date	Location	Number of Committee Members	Fairness Committee Member Demographic Summary	Number of Items Reviewed	Number of Items Rejected by Committees
				Teaching Experience: 20% 1 to 5 years, 15% 6 to 10 years, 10% 11 to 15 years, 25% 16 to 20 years, 30% More than 20 years		
South Dakota	October '19	Pierre, SD	26	**Gender:** 81% Female, 19% Male **Ethnicity:** 4% American Indian or Alaska Native, 4% Asian, 92% White or Caucasian **Region:** 65% Rural, 15% Suburban, 15% Urban, 4% Not Applicable **Teaching Experience:** 12% 1 to 5 years, 12% 6 to 10 years, 19% 11 to 15 years, 19% 16 to 20 years, 38% More than 20 years	[a]	[a]
U.S. Virgin Islands	October '21	Virtual	18	**Gender:** 72% Female, 28% Male **Ethnicity**: 6% Asian, 88% Black or African American, 6% White or Caucasian **Region:** 17% Rural, 17% Urban, 11% Suburban, 17% Not Applicable, 38% Did Not Respond **Teaching Experience:** 22% 1 to 5 years, 5% 6 to 10 years, 17% 11 to 15 years, 17% 16 to 20 years, 39% More than 20 years	299	28
Utah	July '17	Park City, UT	6	**Gender:** 100% Female **Ethnicity:** 33% American Indian or Alaska Native, 33% Hispanic or Latino, 33% White or Caucasian **Region:** 17% Rural, 83% Did Not Respond **Teaching Experience:** 17% General Education, 17% Special Education, 33% Administration, 33% Other	44	2
	December '17	Salt Lake City, UT	6	**Gender:** 83% Female, 17% Make **Ethnicity:** 33% Black or African American, 17% Hispanic or Latino, 33% Native American, 17% Not Applicable **Teaching Experience:** 33% Administration, 83% Other	48	1
	October '19	Provo, UT	11	**Gender:** 27% Female, 73% Did Not Respond **Ethnicity**: 9% Hispanic or Latino, 18% White or Caucasian, 73% Did Not Respond **Region:** 9% Urban, 91% Did Not Respond **Teaching Experience:** 9% 1 to 5 years, 18% More than 20 years, 73% Did Not Respond	31	0
	July '20[b]	Virtual	9	**Gender:** 22% Female, 78% Did Not Respond **Ethnicity**: 22% Hispanic or Latino, 78% Did Not Respond **Region:** 100% Did Not Respond	38	1

State/Item Bank	Date	Location	Number of Committee Members	Fairness Committee Member Demographic Summary	Number of Items Reviewed	Number of Items Rejected by Committees
				Teaching Experience: 11% None, 11% 6 to 10 years, 78% Did Not Respond		
	December '20	Virtual	6	**Gender:** 50% Female, 50% Did Not Respond **Ethnicity**: 17% Hispanic or Latino, 33% White or Caucasian, 50% Did Not Respond **Region:** 17% Suburban, 83% Did Not Respond **Teaching Experience:** 17% 1 to 5 years, 33% 16 to 20 years, 50% Did Not Respond	14	0
	July '21[b]	Virtual	11	**State:** 22% Connecticut, 4% Hawaii, 9% Idaho, 4% Montana, 4% Oregon, 4% Rhode Island, 13% South Dakota, 4% Utah, 3% Vermont, 12% West Virginia, 13% Wyoming, 4% Did Not Respond **Teaching Experience:** 9% 1 to 5 years, 18% 6 to 10 years, 29% 11 to 15 years, 19% 16 to 20 years, 25% More than 20 years	64	0
	August '21	Virtual	6	**Gender:** 100% Female **Ethnicity:** 17% Hispanic or Latino, Native American or Alaskan American, White or Caucasian, Multiracial or Biracial, 83% White or Caucasian **Region:** 33% Rural, 33% Suburban, 17% Urban, 17% Did Not Respond **Teaching Experience:** 17% Less than 1 year, 33% 1 to 5 years, 50% More than 20 years	67	1
	September '22	Salt Lake City, UT	28	**Gender:** 81% Female, 19% Male **Ethnicity:** 2% American Indian or Alaska Native, 1% Asian, 2% Hispanic or Latino, 1% Mediterranean, 80% White or Caucasian **Region:** 26% Rural, 52% Suburban, 16% Urban, 7% Not Applicable **Teaching Experience:** 29% 1 to 5 years, 39% 6 to 10 years, 3% 11 to 15 years, 7% 16 to 20 years, 7% More than 20 years, 16% Not Applicable	111	12
West Virginia	January '17	Charleston, WV	28[c]	Not Collected	34	[a]
	January '19	Charleston, WV	10	**Gender:** 89% Female, 11% Male **Ethnicity:** 11% Black or African American, 89% White or Caucasian **Region:** 100% Rural **Teaching Experience:** 100% General Education	191	[a]

State/Item Bank	Date	Location	Number of Committee Members	Fairness Committee Member Demographic Summary	Number of Items Reviewed	Number of Items Rejected by Committees
	July '21[b]	Virtual	2	**State:** 22% Connecticut, 4% Hawaii, 9% Idaho, 4% Montana, 4% Oregon, 4% Rhode Island, 13% South Dakota, 4% Utah, 3% Vermont, 12% West Virginia, 13% Wyoming, 4% Did Not Respond **Teaching Experience:** 9% 1 to 5 years, 18% 6 to 10 years, 29% 11 to 15 years, 19% 16 to 20 years, 25% More than 20 years	12	1
	June/ August '22[b]	Virtual	3	**State:** 33% Hawaii, 67% Utah **Gender:** 100% Female **Ethnicity:** 33% Asian or Pacific Islander, 33% White or Caucasian, 33% Multiracial or Biracial **Region:** 33% Suburban, 67% Did Not Respond **Teaching Experience:** 33% None, 33% 11 to 15 years, 33% More than 20 years	13	0
Wyoming	December '17	Cheyenne, WY	5	Not Collected	32	3
	October '18	Cheyenne, WY	5	Not Collected	39	0
	November '19	Cheyenne, WY	7	**Gender:** 14% Female, 86% Male **Teaching Experience:** 14% 6 to 10 years, 57% 11 to 20 years, 29% 21 or more years	44	1
	August '20	Virtual	14	**Gender:** 29% Female, 7% Male, 64% Did Not Respond **Ethnicity**: 36% White or Caucasian, 64% Did Not Respond **Region:** 22% Rural, 78% Did Not Respond **Teaching Experience:** 7% 11 to 15 years, 7% 16 to 20 years, 22% More than 20 years, 64% Did Not Respond	37	1
	June/ July '21	Virtual	6	**Gender:** 67% Female, 17% Male, 17% Did Not Respond **Ethnicity:** 83% White or Caucasian, 17% Did Not Respond **Region:** 50% Rural, 17% Suburban, 33% Did Not Respond **Teaching Experience:** 17% 6 to 10 years, 50% 11 to 15 years, 17% More than 20 years, 17% Did Not Respond	39	39
	July '21[b]	Virtual	4	**State:** 22% Connecticut, 4% Hawaii, 9% Idaho, 4% Montana, 4% Oregon, 4% Rhode Island, 13% South Dakota, 4% Utah, 3% Vermont, 12% West Virginia, 13% Wyoming, 4% Did Not Respond **Teaching Experience:** 9% 1 to 5 years, 18% 6 to 10 years, 29% 11 to 15 years, 19% 16 to 20 years, 25% More than 20 years	28	0

State/Item Bank	Date	Location	Number of Committee Members	Fairness Committee Member Demographic Summary	Number of Items Reviewed	Number of Items Rejected by Committees
	June/ August '22[b]	Virtual	6	**State:** 17% Connecticut, 17% Hawaii, 50% Utah, 17% West Virginia **Gender:** 100% Female **Ethnicity:** 17% Asian or Pacific Islander, 50% White or Caucasian, 33% Multiracial or Biracial **Region:** 17% Suburban, 17% Urban, 67% Did Not Respond **Teaching Experience:** 33% None, 33% 11 to 15 years, 33% More than 20 years	37	0

[a]As of the time of writing this report, the number of science items reviewed and/or rejected by Fairness Committees is not currently available.
[b]Items were reviewed in a combined Fairness Committee Meeting that included ICCR and MOU state-owned items. Items reviewed in the combined meetings are displayed by their respective state or bank of ownership.
[c]The number of Committee Members includes the total members across ELA, math, and science committees. The specific number of science committee members is currently unavailable.

Appendix 2-I

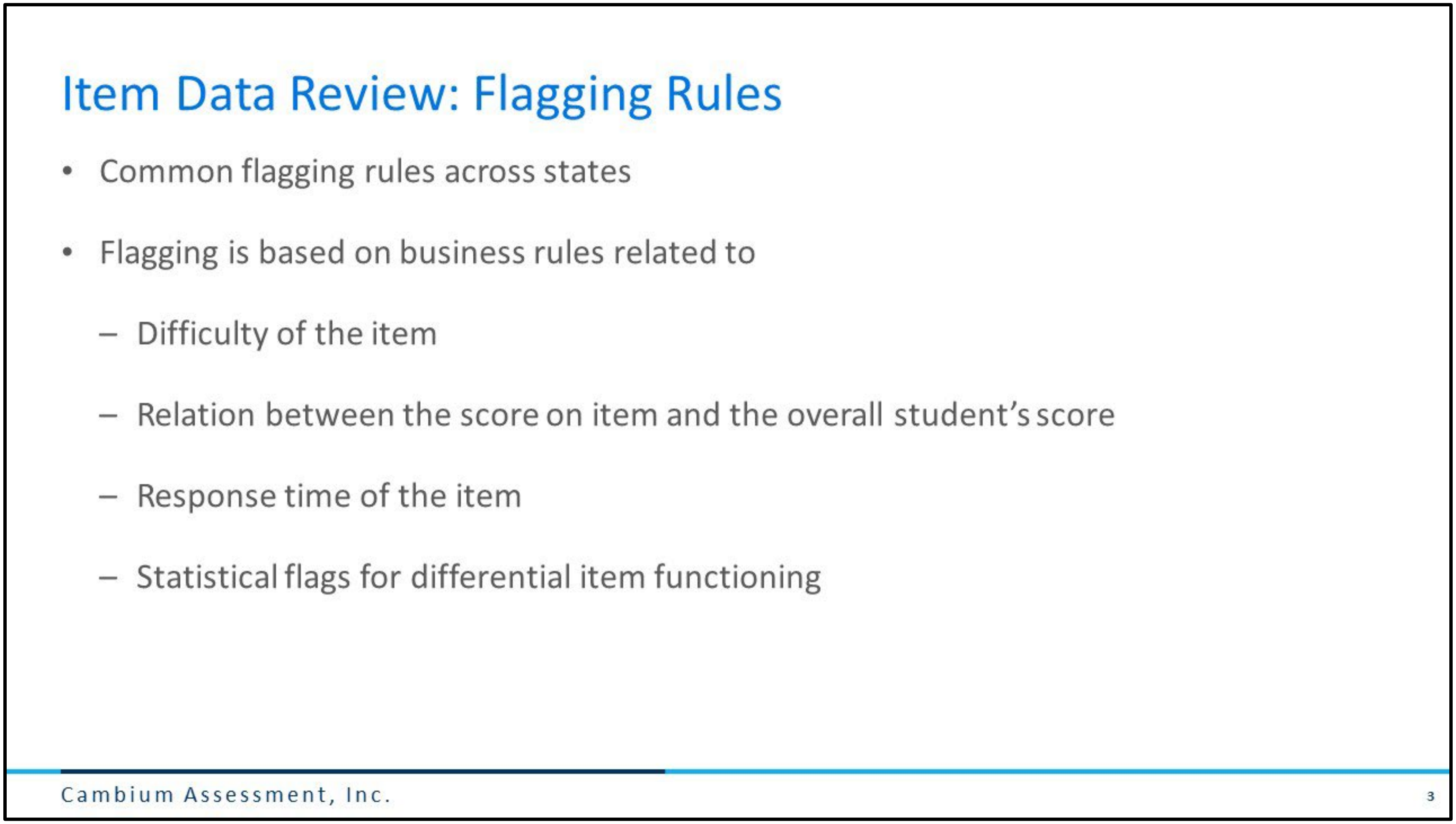
Sample Data Review Training Materials

Sample Data Review Training Materials

Item Data Review

- Item Data Review is the final step before items move to the operational pool
- For every state: data review is carried out for the items owned by the state
- Decision to send an item to data review is based on empirical data
- Statistics are computed at the assertion-level
- Inclusion in data review is decided at the item level, not at the assertion level
  - Inclusion in data review is based on statistical flags that rely on assertion level statistics but are evaluated for the entire item

Cambium Assessment, Inc. 2


Item Data Review: Flagging Rules
• Common flagging rules across states
• Flagging is based on business rules related to
– Difficulty of the item
– Relation between the score on item and the overall student's score
– Response time of the item
– Statistical flags for differential item functioning
Cambium Assessment, Inc.
3

Flagging Rules: *P*-value

- The *p*-value is the proportion of students for which the assertion is TRUE
- Corresponds to the difficulty of an item in a traditional assessment
- Across an item bank, we want to see assertions with *p*-values across the full range to be able to precisely measure proficiency across all proficiency levels
  - A low *p*-value is not bad per se
- However, we want to make sure the low *p*-value is not a result of an item being misleading

Cambium Assessment, Inc. 4

Flagging Rules: *P*-value

- Criteria for clusters:
  - average *p*-value < .30 (across the assertions within a cluster)
  - average *p*-value > .85 (across the assertions within a cluster)
- Criteria for stand-alone items (typically has 1-3 assertions):
  - average *p*-value < . 15 (across the assertions within a stand-alone item)
  - average *p*-value > .95 (across the assertions within a stand-alone item)

Cambium Assessment, Inc. 5

Flagging Rules: Item-Total Correlation

- We expect students who do well on the test overall to have a higher probability of doing well on individual assertions
- The item-total correlation describes that relation
- Criteria
  - Average item-total (biserial) correlation < .25
  - One or more assertions with an item-total correlation < 0.05

Cambium Assessment, Inc. 6

Flagging Rules: Differential Item Functioning

- Fair items behave similar across groups
- Probability of answering correctly is the same for all students of similar ability regardless of group membership
- Groups are defined by
  - Gender
  - Ethnicity
  - Economically disadvantaged vs. not
  - ELL vs. not ELL
  - Special Education vs. not

Cambium Assessment, Inc. 7

Flagging Rules: Differential Item Functioning

- Method: Mantel-Haenszel (Holland & Thayer, 1988)
  - Compares performance on an item (i.e., assertion for assertion-based scoring) between reference and focal groups conditional on overall performance
  - Using the theta estimate as stratification variable

Cambium Assessment, Inc. 8

Flagging Rules: Differential Item Functioning

- Severity of possible bias based on significance testing and effect size
  - "A" No statistical evidence of DIF
  - "B" Evidence for potential mild DIF
  - "C" Evidence for potential severe DIF
- Direction of possible bias
  - "–" assertion favors reference groups (e.g., whites/male/non ELLs)
  - "+" assertion favors focal group

Cambium Assessment, Inc. 9

Flagging Rules: Differential Item Functioning

- DIF Criterion
  - For clusters: 2 or more assertions show ‘C’ DIF in the same direction
  - For stand-alone items: 1 or more assertions show ‘C’ DIF in the same direction

Cambium Assessment, Inc. 10

Flagging Rules: Timing

- We want a good balance between the amount of information an item provides, and the time students spend on the item
- Criteria
  - For clusters: percentile 80 > 15 minutes
    - » A percentile 80 of *x* minutes: 80% of the students spent *x* minutes or less on the item
  - For stand-alone items: percentile 80 > 3 minutes
  - Assertions per minute < .5 for clusters and stand-alone items

Cambium Assessment, Inc. 11

Item Data Review: Process

- Facilitator presents the item
- Item is presented with information on
  - Grade
  - Discipline
  - Disciplinary Core Idea
  - Performance Expectation
- Statistics on the assertions of the item are presented
  - Including the reason for flagging

Cambium Assessment, Inc. 12

Item Data Review: Process

- Evaluation of item (stimulus, interactions, assertions)
- For every item, one of the following decisions is made
  - Reject
  - Accept as is

Cambium Assessment, Inc. 13

Appendix 2-J

Data Review Committee Participant Details

Data Review Committee Participant Details

Table I-1. Data Review Committee Participants, Science

.

State/Item Bank	Date	Location	Number of Committee Members	IDR Committee Member Demographic Summary	Number of Items Reviewed	Number of Items Rejected by Committees
Connecticut	August '18	New Britain, CT	29	**Gender:** 88% Female, 12% Male	18	11
	August '19	Cromwell, CT	29	**Gender:** 83% Female, 17% Male	53	17
	August '21	Virtual	19	**Gender:** 63% Female, 21% Male, 16% Did Not Respond **Ethnicity:** 84% White or Caucasian, 16% Did Not Respond **Region:** 21% Suburban, 21% Urban, 58% Did Not Respond **Teaching Experience:** 5% 1 to 5 years, 5% 6 to 10 years, 16% 11 to 15 years, 21% 16 to 20 years, 37% More than 20 years, 16% Did Not Respond	51	12
	August '22	Virtual	15	**Gender:** 73% Female, 20% Male, 7% Did Not Respond **Ethnicity:** 7% Hispanic or Latino, 87% White or Caucasian, 7% Did Not Respond **Region:** 60% Suburban, 7% Urban, 33% Did Not Respond **Teaching Experience:** 13% 6 to 10 years, 27% 16 to 20 years, 53% More than 20 years, 7% Did Not Respond	19	6
	September '23	Virtual	12	**Gender:** 83% Female, 17% Male **Ethnicity:** 100% White or Caucasian **Region:** 8% Rural, 67% Suburban, 17% Urban, 8% Not Applicable **Teaching Experience:** 17% 11 to 15 years, 25% 16 to 20 years, 58% More than 20 years	14	21
Hawaii	August '18	Honolulu, HI	18	**Gender:** 72% Female, 28% Male	32	3
	August '19	Honolulu, HI	18	**Gender:** 71% Female, 29% Male **Ethnicity:** 12% American Indian and White, 41% Asian, 6% Asian and White, 12% Hispanic and White, 18% Native Hawaiian or Pacific Islander, 12% White or Caucasian **Teaching Experience:** 53% General Education, 6% General Education with SPED Certification, 12% Special Education, 29% Other	37	13

State/Item Bank	Date	Location	Number of Committee Members	IDR Committee Member Demographic Summary	Number of Items Reviewed	Number of Items Rejected by Committees
	August '21[a]	Virtual	7	**State:** 14% Connecticut, 29% Hawaii, 14% Idaho, 29% West Virginia, 14% Wyoming **Gender:** 86% Female, 14% Male **Ethnicity**: 86% White or Caucasian, 14% Did Not Respond **Region:** 14% Rural, 29% Suburban, 57% Did Not Respond **Teaching Experience:** 29% 11 to 15 years, 14% 16 to 20 years, 29% More than 20 years, 14% Did Not Respond	26	8
	August '22[a]	Virtual	12	**State:** 17% Connecticut, 17% Hawaii, 8% Idaho, 25% Oregon, 33% Wyoming **Gender:** 75% Female, 25% Male **Ethnicity:** 8% Asian or Pacific Islander, 82% White or Caucasian **Region:** 50% Rural, 42% Suburban, 8% Did Not Respond **Island:** 8% Not Applicable, 8% Oahu, 75% Did Not Respond **Teaching Experience:** 33% 6 to 10 years, 8% 16 to 20 years, 58% More than 20 years	49	8
	August '23[a]	Virtual	15	**State:** 7% Connecticut, 13% Hawaii, 20% Montana, 7% New Hampshire, 40% Oregon, 7% Utah, 7% Wyoming **Gender:** 60% Female, 40% Male **Ethnicity:** 13% Asian or Pacific Islander, 7% Hispanic or Latino, 7% Native American or Alaskan, 73% White or Caucasian **Region:** 27% Rural, 13% Urban, 60% Did Not Respond **Teaching Experience:** 20% 1 to 5 years, 27% 6 to 10 years, 7% 11 to 15 years, 47% 16 to 20 years	26	5
ICCR	July '18	Virtual	18	Not Collected	84	8
	August '19[b]	Virtual	–	–	43	3
	August '21[a]	Virtual	11	**State:** 27% Connecticut, 9% Hawaii, 18% Idaho, 36% West Virginia, 9% Wyoming **Gender:** 82% Female, 18% Male **Ethnicity:** 54% White or Caucasian, 46% Did Not Respond **Region:** 9% Rural, 27% Suburban, 64% Did Not Respond **Teaching Experience:** 9% 6 to 10 years, 9% 11 to 15 years, 36% More than 20 years, 46% Did Not Respond	75	6
	August '22[a]	Virtual	20	**State:** 15% Connecticut, 20% Idaho, 5% North Dakota, 35% Oregon, 5% South Dakota, 20% Wyoming **Gender:** 85% Female, 15% Male	68	14

State/Item Bank	Date	Location	Number of Committee Members	IDR Committee Member Demographic Summary	Number of Items Reviewed	Number of Items Rejected by Committees
				Ethnicity: 5% Asian or Pacific Islander, 95% White or Caucasian **Region:** 30% Rural, 25% Suburban, 15% Urban, 30% Did Not Respond **Teaching Experience:** 10% 1 to 5 years, 35% 6 to 10 years, 15% 16 to 20 years, 40% More than 20 years		
	August ‘23[a]	Virtual	19	**State:** 11% Connecticut, 21% Hawaii, 5% Montana, 5% New Hampshire, 21% Oregon, 16% Utah, 5% West Virginia, 16% Wyoming **Gender:** 74% Female, 26% Male **Ethnicity:** 16% Asian or Pacific Islander, 84% White or Caucasian **Region:** 21% Rural, 11% Suburban, 5% Urban, 63% Did Not Respond **Teaching Experience:** 21% 1 to 5 years, 16% 6 to 10 years, 16% 16 to 20 years, 47% More than 20 years	54	9
Idaho	August ‘19	[c]	10	**Gender:** 70% Female, 20% Male, 1% Did Not Respond **Ethnicity:** 100% White or Caucasian **Region:** 60% Rural, 40% Suburban **Teaching Experience:** 60% General Education, 2% Administration, 2% Coach	12	6
	August ‘21[a]	Virtual	9	**State:** 11% Hawaii, 56% Idaho, 11% West Virginia, 22% Wyoming **Gender:** 89% Female, 11% Male **Ethnicity:** 89% White or Caucasian, 11% Did Not Respond **Region:** 11% Rural, 22% Suburban, 67% Did Not Respond **Teaching Experience:** 22% 6 to 10 years, 22% 11 to 15 years, 11% 16 to 20 years, 33% More than 20 years, 11% Did Not Respond	60	5
	August ‘22[a]	Virtual	8	**State:** 25% Connecticut, 13% Idaho, 25% Oregon, 38% Wyoming **Gender:** 63% Female, 38% Male **Ethnicity:** 13% Hispanic or Latino, 88% White or Caucasian **Region:** 38% Rural, 50% Suburban, 13% Did Not Respond **Teaching Experience:** 13% 1 to 5 years, 13% 6 to 10 years, 25% 11 to 15 years, 13% 16 to 20 years, 38% More than 20 years	4	0

State/Item Bank	Date	Location	Number of Committee Members	IDR Committee Member Demographic Summary	Number of Items Reviewed	Number of Items Rejected by Committees
	August '23[a]	Virtual	17	**State:** 12% Connecticut, 12% Hawaii, 6% Idaho, 6% Montana, 6% New Hampshire, 24% Oregon, 35% Utah **Gender:** 71% Female, 24% Male, 6% Prefer Not to Say **Ethnicity:** 6% Asian or Pacific Islander, 6% Hispanic or Latino, 6% Native American or Alaskan, 71% White or Caucasian, 12% Did Not Respond **Region:** 29% Rural, 18% Suburban, 53% Did Not Respond **Teaching Experience:** 12% 6 to 10 years, 24% 16 to 20 years, 53% More than 20 years, 12% Did Not Respond	5	0
Montana	September '21	Virtual	4	**Gender:** 50% Female, 50% Did Not Respond **Ethnicity:** 50% White or Caucasian, 50% Did Not Respond **Region:** 50% Rural, 50% Did Not Respond **Teaching Experience:** 25% 6 to 10 years, 25% 16 to 20 years, 50% Did Not Respond	17	4
	September '22	Virtual	5	**Gender:** 100% Female **Ethnicity:** 100% White or Caucasian **Region:** 60% Rural, 40% Suburban **Teaching Experience:** 40% 6 to 10 years, 20% 16 to 20 years, 40% More than 20 years	17	3
	August '23[a]	Virtual	11	**State:** 18% Hawaii, 18% Montana, 9% New Hampshire, 9% Oregon, 36% Utah, 9% West Virginia **Gender:** 100% Female **Ethnicity:** 18% Hispanic or Latino, 73% White or Caucasian, 9% Did Not Respond **Region:** 36% Rural, 64% Did Not Respond **Teaching Experience:** 9% 6 to 10 years, 18% 16 to 20 years, 64% More than 20 years, 9% Did Not Respond	12	3
Multi-State Science Assessment (Rhode Island and Vermont)	August '18	Virtual	–	–	9	6
	August '19	Virtual	–	–	14	4
	August '21	Virtual	–	–	18	9
	September '22	Virtual	–	–	11	7
Oregon	September '18	Salem, OR	11	**Gender:** 82% Female, 18% Male **Ethnicity:** 100% White or Caucasian **Region:** 73% Rural, 27% Urban	44	6

State/Item Bank	Date	Location	Number of Committee Members	IDR Committee Member Demographic Summary	Number of Items Reviewed	Number of Items Rejected by Committees
				Teaching Experience: 64% General Education, 55% Bilingual Education, 36% Special Education, 18% Administration, 18% Other		
	August '19	Virtual	4	**Gender:** 50% Female, 50% Male **Ethnicity:** 100% White or Caucasian **Region:** 50% Rural, 50% Urban **Teaching Experience:** 50% General Education, 25% Bilingual Education, 25% Special Education, 25% Administration, 75% Other	8	7
	August '22[a]	Virtual	8	**State:** 38% Connecticut, 38% Idaho, 13% Wyoming, 13% Did Not Respond **Gender:** 75% Female, 13% Male, 13% Did Not Respond **Ethnicity:** 88% White or Caucasian, 13% Did Not Respond **Region:** 25% Rural, 13% Suburban, 25% Urban, 38% Did Not Respond **Teaching Experience:** 25% 6 to 10 years, 13% 11 to 15 years, 13% 16 to 20 years, 38% More than 20 years, 13% Did Not Respond	31	8
	August '23[a]	Virtual	16	**State:** 6% Connecticut, 6% Idaho, 13% Montana, 38% Oregon, 19% Utah, 6% West Virginia, 13% Wyoming **Gender:** 75% Female, 19% Male, 6% Prefer Not to Say **Ethnicity:** 13% Asian or Pacific Islander, 6% Native American or Alaskan, 75% White or Caucasian, 6% Did Not Respond **Region:** 13% Rural, 19% Suburban, 6% Urban, 63% Did Not Respond **Teaching Experience:** 19% 1 to 5 years, 25% 6 to 10 years, 25% 16 to 20 years, 25% More than 20 years, 6% Did Not Respond	12	2
Rhode Island	September '23	Virtual	–	–	17	6
South Dakota	September '21	Virtual	–	–	15	0
	September '22	Virtual	–	–	4	1
	September '23	Virtual	–	–	6	2
Utah	August '18	Salt Lake City, UT	16	**Gender:** 93% Female, 7% Male **Ethnicity:** 87% White or Caucasian, 13% Did Not Respond **Region:** 13% Suburban, 27% Rural, 60% Did Not Respond **Teaching Experience:** 100% General Education	40	6
	September '21	Virtual	6	**Gender:** 63% Female, 38% Male	11	3

State/Item Bank	Date	Location	Number of Committee Members	IDR Committee Member Demographic Summary	Number of Items Reviewed	Number of Items Rejected by Committees
				Ethnicity: 13% Native Hawaiian or Pacific Islander, 88% White or Caucasian **Region:** 50% Rural, 13% Suburban, 38% Urban **Teaching Experience:** 38% 6 to 10 years, 38% 11 to 15 years, 25% More than 20 years		
	September '22	Salt Lake City, UT	17	**Gender:** 88% Female, 13% Male **Ethnicity:** 6% Asian, 12% Hispanic or Latino, 6% Mixed, 77% White or Caucasian **Region:** 12% Rural, 41% Suburban, 41% Urban, 6% Not Applicable **Teaching Experience:** 6% Less than 1 year, 29% 1 to 5 years, 36% 6 to 10 years, 18% 11 to 15 years, 12% More than 20 years	11	6
	September '23	Salt Lake City, UT	20	**Gender:** 70% Female, 30% Male **Ethnicity:** 5% Asian, 10% Hispanic, 10% Latino, 75% White or Caucasian **Region:** 10% Rural, 50% Suburban, 25% Urban, 15% Not Applicable **Teaching Experience:** 10% Less than 1 year, 25% 1 to 5 years, 20% 6 to 10 years, 20% 11 to 15 years, 10% More than 20 years, 10% Not Applicable, 5% Did Not Respond	6	0
West Virginia	July '18	[c]	4	Not Collected	3	1
	September '19	[c]	4	Not Collected	7	6
	August '21[a]	Virtual	4	**State:** 25% Hawaii, 50% West Virginia, 25% Wyoming **Gender:** 100% Female **Ethnicity**: 75% White or Caucasian, 25% Did Not Respond **Region:** 25% Rural, 25% Suburban, 50% Did Not Respond **Teaching Experience:** 50% 11 to 15 years, 25% More than 20 years, 25% Did Not Respond	7	3
	August '22[a]	Virtual	9	**State:** 22% Connecticut, 33% Idaho, 11% Oregon, 33% Wyoming **Gender:** 89% Female, 11% Male **Ethnicity:** 100% White or Caucasian **Region:** 56% Rural, 11% Suburban, 11% Urban, 22% Did Not Respond **Teaching Experience:** 11% 1 to 5 years, 22% 6 to 10 years, 11% 11 to 15 years, 11% 16 to 20 years, 44% More than 20 years	10	4

State/Item Bank	Date	Location	Number of Committee Members	IDR Committee Member Demographic Summary	Number of Items Reviewed	Number of Items Rejected by Committees
	August ‘23[a]	Virtual	11	**State:** 9% Connecticut, 18% Hawaii, 9% Montana, 9% New Hampshire, 9% Oregon, 45% Utah **Gender:** 82% Female, 9% Male, 9% Prefer Not to Say **Ethnicity:** 9% Hispanic or Latino, 82% White or Caucasian, 9% Did Not Respond **Region:** 36% Rural, 9% Suburban, 56% Did Not Respond **Teaching Experience:** 9% 6 to 10 years, 27% 16 to 20 years, 56% More than 20 years, 9% Did Not Respond	12	3
Wyoming	October ‘18	Cheyenne, WY	12	**Gender:** 75% Female, 25% Male **Teaching Experience:** 8% 3 to 5 years, 8% 6 to 10 years, 58% 11 to 20 years, 25% 21 or more years	16	6
	August ‘19	Cheyenne, WY	10	**Gender:** 90% Female, 10% Male **Region:** 40% Suburban, 60% Rural **Teaching Experience:** 90% General Education, 10% Administration	16	5
	August ‘21[a]	Virtual	8	**State:** 38% Connecticut, 13% Hawaii, 13% West Virginia, 38% Wyoming **Gender:** 75% Female, 25% Male **Ethnicity**: 75% White or Caucasian, 25% Did Not Respond **Region:** 13% Rural, 25% Suburban, 13% Urban, 50% Did Not Respond **Teaching Experience:** 13% 11 to 15 years, 63% More than 20 years, 25% Did Not Respond	16	4
	August ‘22[a]	Virtual	12	**State:** 17% Connecticut, 8% Hawaii, 17% Idaho, 17% Oregon, 42% Wyoming **Gender:** 67% Female, 33% Male **Ethnicity:** 8% Asian or Pacific Islander, 8% Hispanic or Latino, 83% White or Caucasian **Region:** 42% Rural, 50% Suburban, 8% Did Not Respond **Teaching Experience:** 8% 1 to 5 years, 25% 6 to 10 years, 17% 11 to 15 years, 8% 16 to 20 years, 42% More than 20 years	19	3
	August ‘23[a]	Virtual	17	**State:** 12% Connecticut, 12% Hawaii, 6% Idaho, 6% Montana, 6% New Hampshire, 24% Oregon, 35% Utah **Gender:** 71% Female, 24% Male, 6% Prefer Not to Say **Ethnicity:** 6% Asian or Pacific Islander, 6% Hispanic or Latino, 6% Native American or Alaskan, 71% White or Caucasian, 12% Did Not Respond	8	1

State/Item Bank	Date	Location	Number of Committee Members	IDR Committee Member Demographic Summary	Number of Items Reviewed	Number of Items Rejected by Committees
				Region: 29% Rural, 18% Suburban, 53% Did Not Respond **Teaching Experience:** 18% 6 to 10 years, 18% 16 to 20 years, 53% More than 20 years, 12% Did Not Respond		

Note. MSSA, Rhode Island, and South Dakota-owned items were reviewed by Rhode Island Department of Education and Vermont Agency of Education science content experts, the Rhode Island Department of Education, and the South Dakota Department of Education, respectively.

[a]Combined Item Data Review Meetings were conducted for multiple states in 2021, 2022, and 2023 (184 items were reviewed in the combined meeting format for Hawaii, Idaho, West Virginia, Wyoming, and ICCR items in 2021; 181 items were reviewed in the combined meeting format for Hawaii, Idaho, Oregon, West Virginia, Wyoming, and ICCR items in 2022, and 129 items were reviewed in the combined meeting format for Hawaii, Idaho, Montana, Oregon, West Virginia, Wyoming, and ICCR items in 2023). In 2021, 25 committee members took part in the combined Item Data Review Meetings; in 2022, 38 committee members participated in the combined Item Data Review Meetings, and in 2023, 41 committee members participated in the combined Item Data Review Meetings. Items reviewed in the combined meetings are displayed by their respective state or bank of ownership.

[b]During the summer 2019, ICCR field-test items underwent committee review in Connecticut, Hawaii, and Idaho.

[c]The specific location of the Data Review Committee Meeting is unavailable at the time of writing this report.

Appendix 2-K

Example Item Interactions

Interaction Types Available in the Montana Science (MSA) Assessment

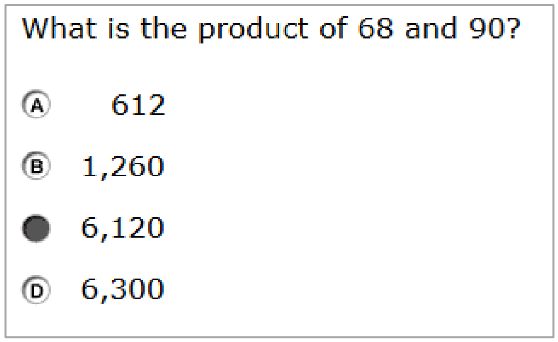
Review of Different Interaction Types

Interaction Type	Associated Sub-Types	Legacy Item Types Supported
Choice	Multiple Choice	MC
	Multiple Select	MS
	Scaffolding	ASI2, ASI3
Text Entry	Simple Text Entry	EA, ECR, LA, OE, SA, SR, WCR, RW, SCR
	Embedded Text Entry	CL, FI
	Natural Language	NL
	Extended Response	ER
Table	Table Match	MI
	Table Input	TI
	Column Match	MI
Edit Task	Edit Task	ET
	Edit Task with Choice	ETC
	Edit Task Inline Choice	ETC
Hot Text	Selectable	HTQ
	Re-orderable	HT
	Drag-from-Palette	DnD
	Custom	HTQ, HT, DnD
Equation	N/A	EQN
Grid	Grid	GI
	Hot Spot	GI
	Graphic Gap Match	GI
Simulation*	N/A	SIM

Note. the abbreviations correlate to the attributes used in CAI's Item Tracking System

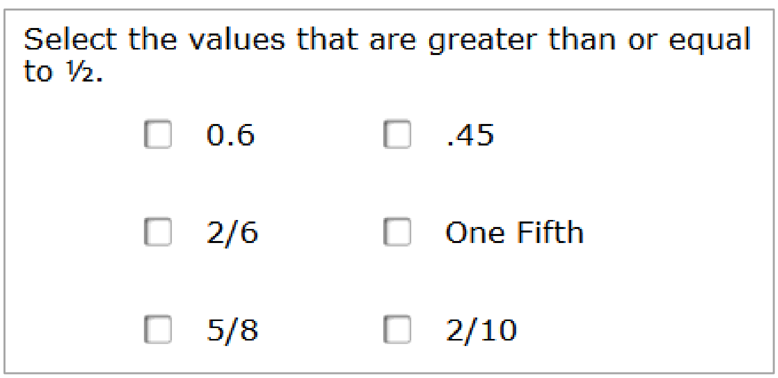
Multiple-Choice Interactions

Multiple-Choice (MC) interactions require students to select a single option from a list of possible answer options. The number and orientation of answer options in a multiple-choice interaction are configurable. Answer options may appear vertically, horizontally, vertically-stacked (in a specified number of columns), or horizontally-stacked (in a specified number of rows).



Multiple-Select Interactions

Multiple-Select interactions require students to select one or more options from a list of possible answer options. The number and orientation of answer options in a multiple-select interaction are configurable. Answer options may appear vertically, horizontally, horizontally-stacked (in a specified number of rows), or vertically-stacked (in a specified number of columns).



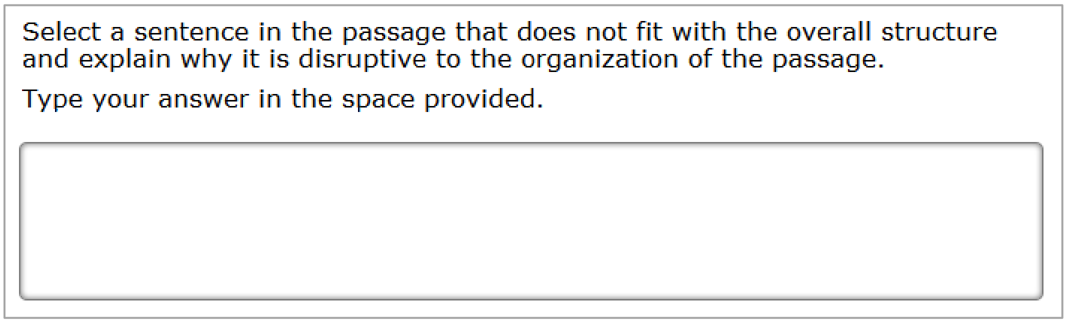
Text Entry Interactions

The Text Entry Interaction Editor allows you to create content for the following interaction types:

- Simple Text Entry Interactions
- Embedded Text Entry Interactions
- Natural Language Interactions
- Extended Response Interactions

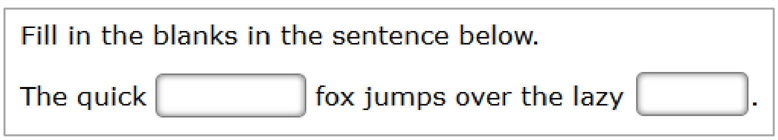
Simple Text Entry Interactions

Simple Text Entry interactions require students to type a response in a text box. For Simple Text Entry interactions, we can allow you to specify the maximum response length for the text box and the type of text editor available to students.



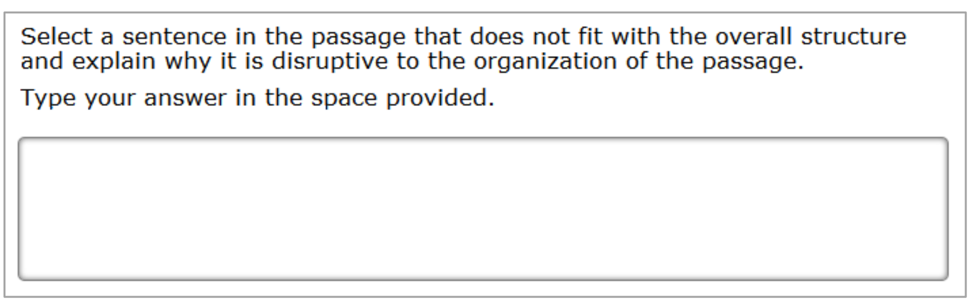
Embedded Text Entry Interactions

Embedded Text Entry interactions require students to type their response in one or more text boxes that are embedded in a section of read-only text.



Extended Response Interactions

Extended Response interactions require students to type a response in a text box. Extended Response interactions are scored by an uploaded essay scoring model that analyzes the student's response to identify variations of acceptable key words and phrases. For Extended Text Entry interactions, we can allow you to specify the maximum response length for the text box and the type of text editor available to students.



Alert: Extended Response interactions cannot be combined with any other interactions in the item.

Table Entry Interaction

The Table Entry Interaction Editor allows you to create content for the following interaction types:

- Table Match Interactions
- Table Input Interactions
- Column Match Interactions

Table Match Interactions

Table Match interactions arrange two sets of match options in a table, with one set listed in columns and the other set listed in rows. Students match options in the columns to options in the rows by marking checkboxes in the cells where the columns and rows intersect.

For each number listed in the rows of the table, mark the checkboxes for each column that describes that number.

	Perfect Square	Prime Number	Odd Number	Even Number
5	☐	☐	☐	☐
12	☐	☐	☐	☐
9	☐	☐	☐	☐

Table Match interactions allow you to customize the number of match options in each set and enter the content for each match option. You can also set restrictions on the number of matches students can make. By default, the panel includes a basic table consisting of three rows and columns (including the row header and column header).

Table Input Interactions

Table Input interactions provide students with a table that includes one or more blank cells. Each blank cell displays a text box in which students can type their response.

Enter a stage direction that you might give to each theater technician listed in the table below.

The first one has been done for you.

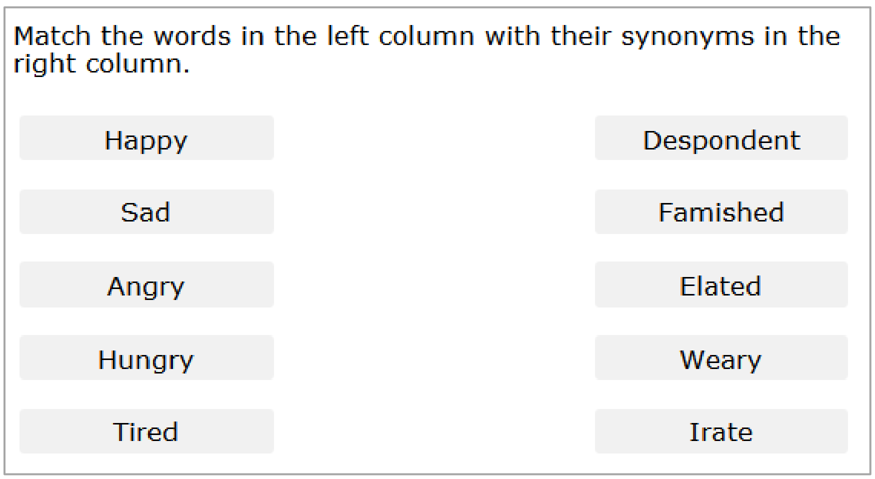
Theater technicians	Stage direction
Set designer	A circular bench around a small obelisk
Props manager	
Sound technician	
Lighting technician	

Table Input interactions allows you to customize the number of rows and columns in the table, specify which cells display text boxes, and enter content for the read-only cells. By default, the panel includes a basic table consisting of three rows and columns (including the row header and column header).

Alert: If a table does not include row headers, then it must include column headers. If a table does not include column headers, then it must include row headers.

Column Match Interactions

Column Match interactions provide students with two columns that each contain a set of match options. Students respond to the interaction by selecting a match option in the left column and then selecting the corresponding match option in the right column. A match option in one set may have one, multiple, or no matches in the other set.



Column Match interactions allows you to customize the number of match options in each set and enter the content for each match option. By default, the panel includes two single-column tables, each of which includes two match options. You can also set restrictions on the number of matches students can make.

Edit Task Interactions

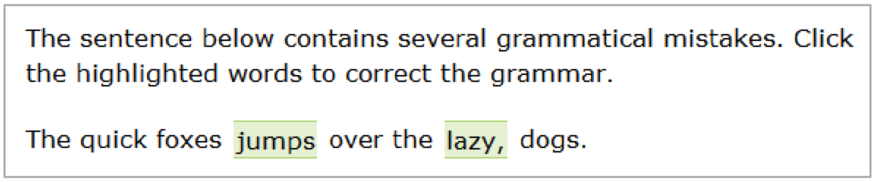
The Edit Task Interaction Editor allows you to create content for the following interaction types:

- Edit Task Interactions
- Edit Task with Choice Interactions
- Edit Task Inline Choice Interactions

Edit Task Interactions

Edit Task interactions provide students with a sentence or paragraph containing one or more tagged text elements. Tagged elements usually contain an error, such as improper spelling or grammar.

To respond to these interactions, students click a tagged element and enter corrected text in an editing window. The entered text replaces the original tagged text.



Edit Task interactions allow you to enter the text that appears in the response area and tag elements within the text that students can edit.

Warning: You cannot include hand-scored and machine-scored interactions in the same item.

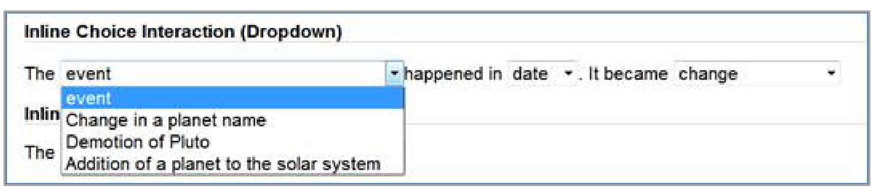
Edit Task with Choice Interactions

Edit Task with Choice interactions are similar to Edit Task interactions. The only difference is that when responding to Edit Task with Choice interactions, students replace the tagged text elements with options selected from a drop-down list.

Edit Task with Choice interactions allow you to enter the text that appears in the response area and tag elements within the text that students can edit.

Edit Task Inline Choice Interactions

Edit Task Inline Choice interactions are similar to Edit Task with Choice interactions. The only difference is that students select replacement options from a drop-down list embedded within the read-only text, rather than accessing the drop-down list via a pop-up window.



Hot Text Interactions

The Hot Text Interaction Editor allows you to create content for the following interaction types:

- Selectable Hot Text Interactions
- Re-orderable Hot Text Interactions
- Drag-from-Palette Hot Text Interactions
- Custom Hot Text Interactions

Selectable Hot Text Interactions

Selectable Hot Text interactions require students to select one or more text elements in the response area.

> Select the sentences that support the inference that the area is in danger of losing its moose population. Select **all** that apply.
>
> A similar boom-and-bust cycle occurs between predator and prey. Ten times the size of a wolf, a moose has long, strong legs and a dangerous kick. So wolves prey mainly on old and weak animals. Good hunting means food for the whole pack. Wolves then raise lots of pups, and their numbers increase. More wolves mean more mouths to feed and more moose get eaten. However, when the moose population decreases, wolves starve.

Selectable Hot Text interactions allows you to set the minimum and maximum number of elements students can select, enter the text that appears in the response area, and tag the text elements that will be selectable.

Re-orderable Hot Text Interactions

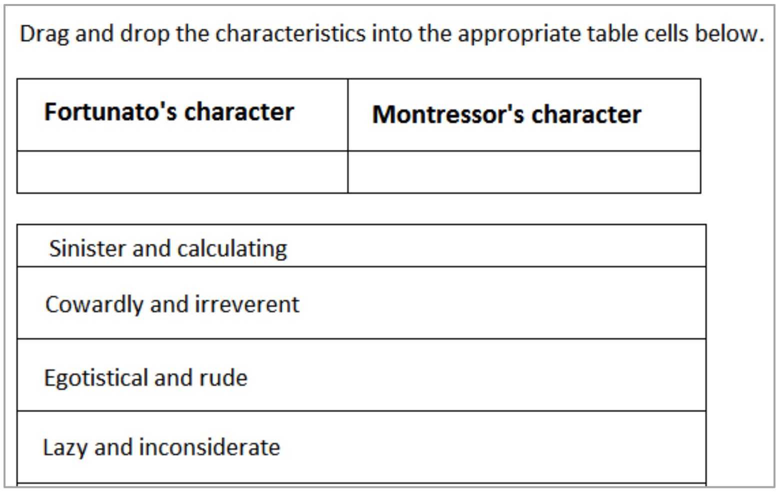
Re-orderable Hot Text interactions require students to click and drag hot text elements into a different order.

> Place the following sentences in the correct order.
>
> Hey Jude. And make it better. Don't be afraid. Take a sad song.

Re-orderable Hot Text interactions allow you to enter the re-orderable text elements in the response area. You can specify the elements' orientation and set them to appear in random order to students.

Drag-from-Palette Hot Text Interactions a.k.a. Hot Text Gap Match

Drag-from-Palette Hot Text interactions require students to drag elements from a palette into the available blank table cells or "gaps" (text boxes) in the response area. Palette elements may consist of text and/or images. Students may be able to drag the same palette element into multiple gaps, depending on the interaction's configuration.



Drag-from-Palette Hot Text interactions allow you to enter the elements that appear in the palette, enter static text for the response area, and create the gap targets where students can drag the text elements. You can enter all of the elements in a single text box or enter each segment in its own text box.

- Can set a minimum/maximum number of times a student is required/allowed to use a specific palette object
- Only supports drag-and-drop of palette items (images or plain text) onto pre-defined drop targets ("gaps" or "blanks") in the body text
    - These palette items are always confined to a special palette region (no "preplacing" them)
    - There is some control over palette placement
    - The items can only be placed in predefined "target" regions

Custom Hot Text Interactions

Custom Hot Text interactions combine the functionality of the other Hot Text interaction sub-types. Students responding to a Custom Hot Text interaction may need to select text elements, rearrange text elements, and/or drag text elements from a palette to blank table cells or drop targets in the response area. In many ways, this is the grid of the text-interaction world. In practice, it is typically used to do drag-and-drop with text, but it can technically do more:

- Supports dragging and dropping text elements onto drop target areas
    - Text elements can originally be placed anywhere in the interaction (there's no dedicated palette)
    - Multiple elements can be dropped onto a target
        - this constitutes a "group"
        - much like grid hotspots, you can set constraints on the group

        - Supports selectable text elements
    - Like grid hotspots, these too can be grouped

Use the word bank to fill in the blank in the sentence below. Then, select all the words in the sentence that are nouns.

Word bank:

young dull good rich

Sentence:

All work and no play makes Jack a ________ boy.

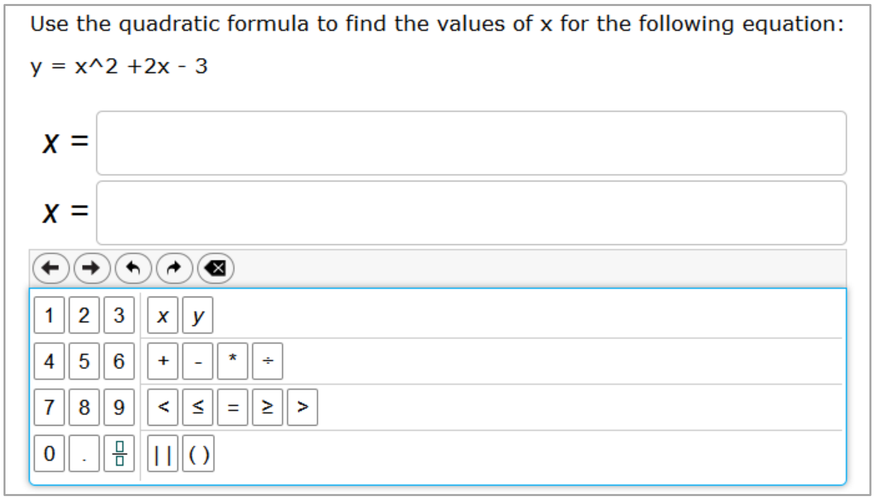
Custom Hot Text interactions allow you to create groups of text elements, as well as the drop targets and static text that appear in the response area. When you create a group of text elements, you must assign a Hot Text functionality to that group. The following functionalities are available:

- **Selectable:** When you assign this functionality to a group, the text elements in the group behave like elements in a Selectable Hot Text interaction. You cannot add drop target elements to this kind of group.
- **Draggable:** When you assign this functionality to a group, the text elements in the group behave like elements in a Re-Orderable Hot Text interaction. If you assign this functionality to a group and also add drop targets to the group, the text elements in the group behave like elements in a Drag-from-Palette Hot Text interaction.

You can create as many groups as you wish, but you can only assign one Hot Text functionality to each group.

Equation Interaction Editor

The Equation Interaction Editor allows you to create content for Equation interactions only. Equation interactions require students to enter a response into input boxes using an on-screen keypad, which may consist of special mathematics characters. Students can also enter their response via a physical keyboard, but they cannot enter any characters that are not included in the on-screen keyboard.



Equation interactions allow you to select the buttons to include in the on-screen keypad, enter static text in the response area, and specify the number of input boxes to include in the response area. When selecting buttons to include in the keypad, you can add individual buttons or an entire row or tab of buttons.

Grid Interactions

The Grid Interaction Editor allows you to create content for the following interaction types:

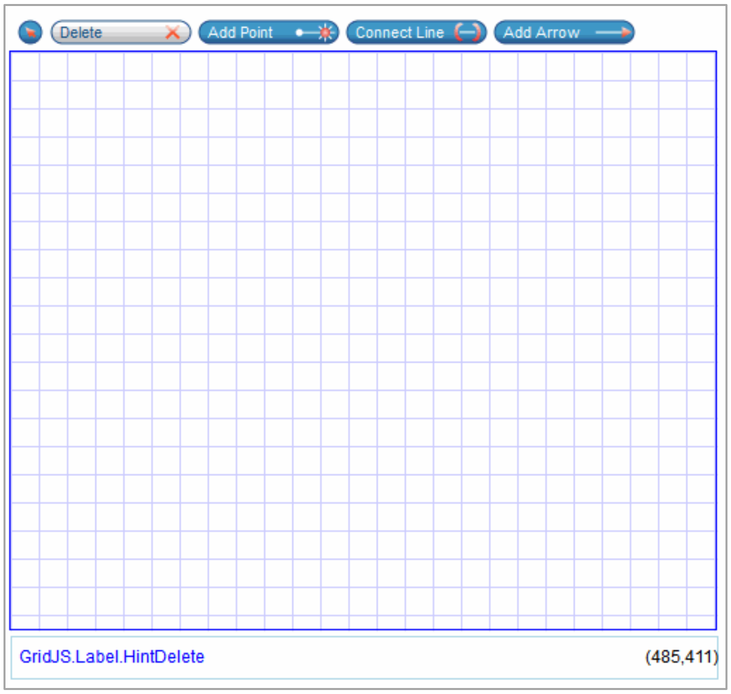
- Grid Interactions
- Hot Spot Interactions
- Graphic Gap Match Interactions

Note: Although there are three options available in the **Interaction Type** drop-down list, the generic **Grid** option allows you to create interactions with functionality similar to Hot Spot and Graphic Gap Match sub-types.

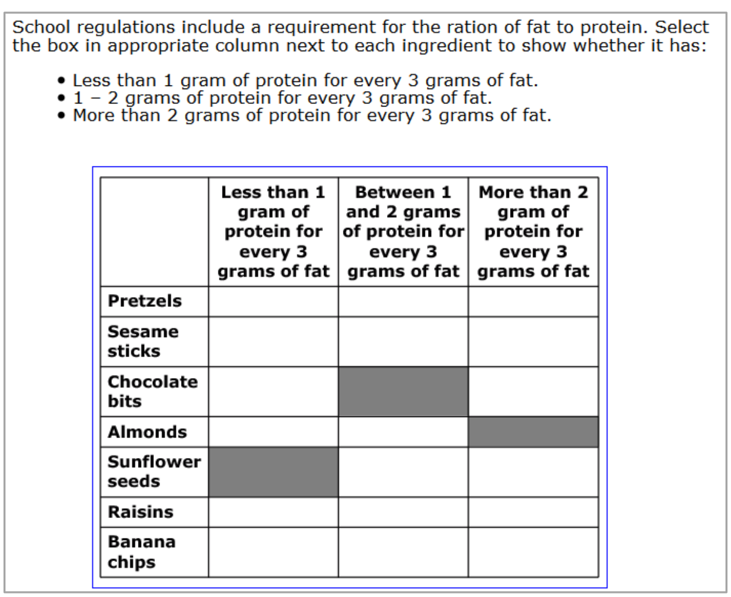
Grid Interactions Types

Grid interactions require students to enter a response by interacting with a grid area in the answer space. There are three general ways in which students can interact with the grid area.

- **Graphing Functionality:** Students can use various tool buttons to add points, lines, and other geometric shapes to the grid area. Only the Grid interaction sub-type allows you to create interactions with this functionality.



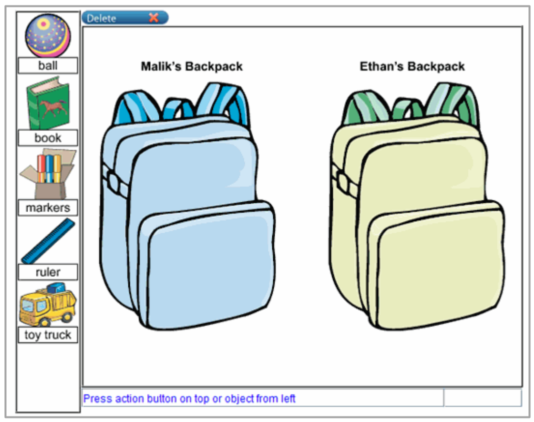
- **Hot Spot Functionality:** Students can click or hover over interactive regions in the grid area (hot spots) in order to activate them. Activated hot spots become highlighted, become outlined, or display an image. The Grid and Hot Spot interaction sub-types allow you to create interactions with this functionality.
    - Hotspots can be defined in groups, each of which can have its own selection constraints
    - These regions support events so clicking a hotspot might change the appearance of the interaction by showing/hiding other images, for example


School regulations include a requirement for the ration of fat to protein. Select the box in appropriate column next to each ingredient to show whether it has:

- Less than 1 gram of protein for every 3 grams of fat.
- 1 – 2 grams of protein for every 3 grams of fat.
- More than 2 grams of protein for every 3 grams of fat.

	Less than 1 gram of protein for every 3 grams of fat	Between 1 and 2 grams of protein for every 3 grams of fat	More than 2 gram of protein for every 3 grams of fat
Pretzels			
Sesame sticks			
Chocolate bits			
Almonds			
Sunflower seeds			
Raisins			
Banana chips			

- **Drag-and-Drop Functionality:** Students can click image or text objects and drag them into various locations in the grid area. The objects for these interactions are either provided in a palette beside the grid area or pre-placed within the grid area itself. The Grid and Graphic Gap Match interaction sub-types allow you to create interactions with this functionality; however, only Graphic Gap Match interactions allow text objects.
  - These palette items can be “preplaced” on the canvas or listed in a separate palette
  - The items can be placed anywhere on the canvas or guided to specific regions with snap points



Note: The functionalities of these interaction types are not mutually exclusive. A single Grid interaction may require students to select hot spots and place objects, or graph lines and select hot spots, and so on. However, a Grid interaction cannot include preplaced objects if it also includes the **Delete** tool button above the grid area.

Grid Hot Spot Interactions

Hot Spot interaction sub-types allow you to create Grid interactions with hot spot functionality. These interactions require students to select hot spot regions in the grid area.

- Only supports click-to-select “hotspots”
  - No visual side-effect events are supported
  - No hotspot groups are supported

Grid Graphic Gap Match Interactions

Graphic Gap Match interactions allow you to create Grid interactions with both hot spot and drag-and-drop functionality. These interactions require students to drag image objects from a palette to hot spot regions (gaps) in the grid area.

- Only supports drag-and-drop of palette items (images or plain text) onto the canvas/background

- These palette items are always confined to a special palette region (no "preplacing" them on the canvas)
- The items can only be placed in predefined "target" regions

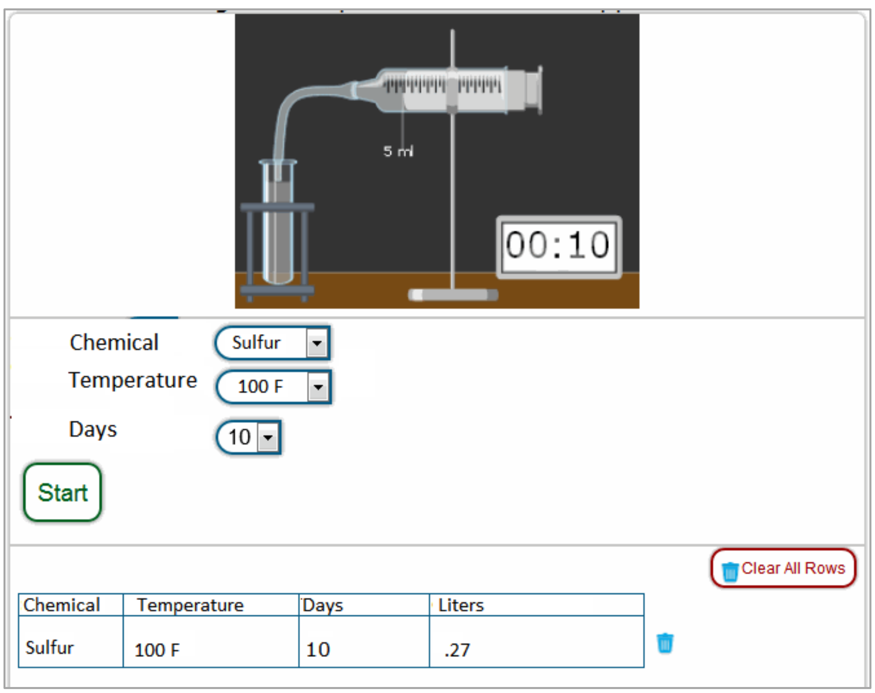
> **Alert:** Graphic Gap Match interactions do not allow you to enable Snap-to-Point or Snap-to-Grid mode. You cannot pre-place image or text objects in the grid area with Graphic Gap Match Interactions.

Basically, graphic gap match and hotspot are dedicated interactions that don't support all the features of a grid. The trade-off here is:

- Graphic gap match and hotspot interactions are rendered differently (more simplistically)
- In some ways, graphic gap match and hotspot are easier to author and maintain
- Grid interactions need to use the "grid rubric tool," which is quite complicated

Simulation Interaction Editor

The Simulation Interaction Editor allows you to create content for Simulation interactions only. Simulation interactions consist of an animation tool, a set of input tools, and an output table. Students select parameters from the input tools to influence the animation. After the animation runs, the simulation results appear in the output table. Students can run multiple trials with different parameters to insert additional rows into this table.



Appendix 2-L

Shared Science Assessment Item Bank

Shared Science Assessment Item Bank

Table K-1. Spring 2023 Shared Science Assessment Operational and Field-Test Item Bank by Performance Expectation, Elementary School

Science Discipline	Disciplinary Core Idea	Performance Expectation	ICCR Items	Montana Items	MOU Items[a]	Total Bank Items
Earth and Space Sciences	ESS1	4-ESS1-1	2	1	13	16
		5-ESS1-1	2	2	13	17
		5-ESS1-2	7	2	13	22
	ESS2	3-ESS2-1	3	1	8	12
		3-ESS2-2	4	1	11	16
		4-ESS2-1	4	1	11	16
		4-ESS2-2	4	0	13	17
		5-ESS2-1	0	1	14	15
		5-ESS2-2	4	0	13	17
	ESS3	3-ESS3-1	4	2	8	14
		4-ESS3-1	7	0	6	13
		4-ESS3-2	9	0	13	22
		5-ESS3-1	4	2	10	16
Life Sciences	LS1	3-LS1-1	7	2	7	16
		4-LS1-1	9	1	12	22
		4-LS1-2	3	1	15	19
		5-LS1-1	3	1	15	19
	LS2	3-LS2-1	4	1	12	17
		5-LS2-1	2	1	16	19
	LS3	3-LS3-1	3	2	11	16
		3-LS3-2	4	0	9	13
	LS4	3-LS4-1	4	1	12	17
		3-LS4-2	9	1	7	17
		3-LS4-3	4	2	8	14
		3-LS4-4	6	1	8	15
Physical Sciences	PS1	5-PS1-1	5	0	13	18
		5-PS1-2	3	0	10	13

Science Discipline	Disciplinary Core Idea	Performance Expectation	ICCR Items	Montana Items	MOU Items[a]	Total Bank Items
		5-PS1-3	4	2	10	16
		5-PS1-4	2	0	12	14
	PS2	3-PS2-1	4	1	8	13
		3-PS2-2	5	2	6	13
		3-PS2-3	5	1	7	13
		3-PS2-4	4	2	7	13
		5-PS2-1	4	1	8	13
	PS3	4-PS3-1	4	0	15	19
		4-PS3-2	3	1	15	19
		4-PS3-3	3	0	12	15
		4-PS3-4	5	1	13	19
		5-PS3-1	4	1	10	15
	PS4	4-PS4-1	2	2	10	14
		4-PS4-2	1	1	13	15
		4-PS4-3	4	0	14	18
Total			**174**	**42**	**461**	**677**

[a]Other MOU states include Connecticut, Hawaii, Idaho, Oregon, Rhode Island, Utah, West Virginia, and Wyoming.

Table K-2. Spring 2023 Shared Science Assessment Operational and Field-Test Item Bank by Performance Expectation, Middle School

Science Discipline	Disciplinary Core Idea	Performance Expectation	ICCR Items	Montana Items	MOU Items[a]	Total Bank Items[b]
Earth and Space Sciences	ESS1	MS-ESS1-1	5	1	9	15
		MS-ESS1-2	2	0	9	11
		MS-ESS1-3	3	2	7	12
		MS-ESS1-4	4	1	11	16
	ESS2	MS-ESS2-1	3	0	7	10
		MS-ESS2-2	4	2	8	14
		MS-ESS2-3	4	0	9	13
		MS-ESS2-4	2	2	6	10
		MS-ESS2-5	3	0	10	13
		MS-ESS2-6	6	0	3	9
	ESS3	MS-ESS3-1	3	2	8	13
		MS-ESS3-2	4	1	8	13
		MS-ESS3-3	2	1	11	14
		MS-ESS3-4	2	1	11	14
		MS-ESS3-5	5	0	10	15
Life Sciences	LS1	MS-LS1-1	2	1	10	13
		MS-LS1-2	2	0	11	13
		MS-LS1-3	2	0	9	11
		MS-LS1-4	5	0	5	10
		MS-LS1-5	2	1	9	12
		MS-LS1-6	4	0	8	12
		MS-LS1-7	4	0	6	10
		MS-LS1-8	6	0	7	13
	LS2	MS-LS2-1	6	3	10	19
		MS-LS2-2	5	0	7	12
		MS-LS2-3	3	1	9	13
		MS-LS2-4	8	0	10	18
		MS-LS2-5	4	1	9	14
	LS3	MS-LS3-1	2	1	8	11

Science Discipline	Disciplinary Core Idea	Performance Expectation	ICCR Items	Montana Items	MOU Items[a]	Total Bank Items[b]
		MS-LS3-2	4	1	9	14
	LS4	MS-LS4-1	5	0	8	13
		MS-LS4-2	2	0	9	11
		MS-LS4-3	3	0	9	12
		MS-LS4-4	3	1	8	12
		MS-LS4-5	5	0	10	15
		MS-LS4-6	2	0	9	11
Physical Sciences	PS1	MS-PS1-1	3	1	7	11
		MS-PS1-2	2	1	8	11
		MS-PS1-3	5	0	7	12
		MS-PS1-4	3	1	10	14
		MS-PS1-5	2	2	9	13
		MS-PS1-6	3	0	5	8
	PS2	MS-PS2-1	2	0	9	11
		MS-PS2-2	1	0	9	10
		MS-PS2-3	2	0	8	10
		MS-PS2-4	2	0	11	13
		MS-PS2-5	1	3	11	15
	PS3	MS-PS3-1	3	1	7	11
		MS-PS3-2	2	0	9	11
		MS-PS3-3	7	0	6	13
		MS-PS3-4	3	0	8	11
		MS-PS3-5	5	0	7	12
	PS4	MS-PS4-1	3	1	7	11
		MS-PS4-2	6	0	8	14
		MS-PS4-3	2	0	10	12
Total			**188**	**33**	**463**	**684**

[a]*Other MOU states include Connecticut, Hawaii, Idaho, Oregon, Rhode Island, Utah, West Virginia, and Wyoming.* [b]*Count excludes eight MOU items that do not align to the NGSS.*

Appendix 2-M

Montana Science Assessment Item Pool

Montana Science Assessment Item Pool

Table L-1. Spring 2023 Montana Science Assessment Operational and Field-Test Item Pool by Performance Expectation, Grade 5

Science Discipline	Disciplinary Core Idea	Performance Expectation	ICCR Items	Montana Items	MOU Items	Total Pool Items
Earth and Space Sciences	ESS1	4-ESS1-1	2	1	0	3
		5-ESS1-1	2	2	6	10
		5-ESS1-2	5	2	4	11
	ESS2	3-ESS2-1	3	1	2	6
		3-ESS2-2	1	1	1	3
		4-ESS2-1	4	1	0	5
		4-ESS2-2	2	0	3	5
		5-ESS2-1	0	1	3	4
		5-ESS2-2	1	0	2	3
	ESS3	3-ESS3-1	3	2	2	7
		4-ESS3-1	4	0	1	5
		4-ESS3-2	4	0	1	5
		5-ESS3-1	3	2	5	10
Life Sciences	LS1	3-LS1-1	1	2	1	4
		4-LS1-1	5	1	1	7
		4-LS1-2	1	1	2	4
		5-LS1-1	3	1	1	5
	LS2	3-LS2-1	4	1	7	12
		5-LS2-1	1	1	2	4
	LS3	3-LS3-1	2	2	0	4
		3-LS3-2	3	0	2	5
	LS4	3-LS4-1	3	1	0	4
		3-LS4-2	4	1	0	5
		3-LS4-3	2	2	1	5
		3-LS4-4	3	1	1	5
Physical Sciences	PS1	5-PS1-1	4	0	1	5

Science Discipline	Disciplinary Core Idea	Performance Expectation	ICCR Items	Montana Items	MOU Items	Total Pool Items
		5-PS1-2	1	0	1	2
		5-PS1-3	4	2	4	9
		5-PS1-4	1	0	3	4
	PS2	3-PS2-1	2	1	2	5
		3-PS2-2	2	2	2	6
		3-PS2-3	3	1	1	5
		3-PS2-4	1	2	2	5
		5-PS2-1	2	1	4	7
	PS3	4-PS3-1	2	0	2	4
		4-PS3-2	1	1	1	3
		4-PS3-3	1	0	2	3
		4-PS3-4	4	1	1	6
		5-PS3-1	3	1	4	8
	PS4	4-PS4-1	2	2	2	6
		4-PS4-2	1	1	1	3
		4-PS4-3	2	0	1	3
Total			**102**	**42**	**81**	**225**

[a]Other MOU states include Connecticut, Hawaii, Idaho, Oregon, Rhode Island, Utah, West Virginia, and Wyoming.

Table L-2. Spring 2023 Montana Science Assessment Operational and Field-Test Item Pool by Performance Expectation, Grade 8

Science Discipline	Disciplinary Core Idea	Performance Expectation	ICCR Items	Montana Items	MOU Items[a]	Total Item Pool
Earth and Space Sciences	ESS1	MS-ESS1-1	2	1	1	4
		MS-ESS1-2	2	0	2	4
		MS-ESS1-3	1	2	1	4
		MS-ESS1-4	1	1	2	4
	ESS2	MS-ESS2-1	1	0	3	4
		MS-ESS2-2	1	2	2	5
		MS-ESS2-3	2	0	1	3
		MS-ESS2-4	1	2	1	4
		MS-ESS2-5	2	0	0	2
		MS-ESS2-6	1	0	1	2
	ESS3	MS-ESS3-1	2	2	1	5
		MS-ESS3-2	2	1	1	4
		MS-ESS3-3	1	1	2	4
		MS-ESS3-4	2	1	6	9
		MS-ESS3-5	3	0	0	3
Life Sciences	LS1	MS-LS1-1	0	1	0	1
		MS-LS1-2	1	0	1	2
		MS-LS1-3	1	0	2	3
		MS-LS1-4	3	0	2	5
		MS-LS1-5	1	1	1	3
		MS-LS1-6	2	0	2	4
		MS-LS1-7	1	0	1	2
	LS2	MS-LS2-1	4	3	4	11
		MS-LS2-2	1	0	0	1
		MS-LS2-3	1	1	1	3
		MS-LS2-5	1	1	1	3
	LS3	MS-LS3-1	0	1	2	3
		MS-LS3-2	2	1	3	6

Science Discipline	Disciplinary Core Idea	Performance Expectation	ICCR Items	Montana Items	MOU Items[a]	Total Item Pool
	LS4	MS-LS4-1	4	0	0	4
		MS-LS4-2	1	0	1	2
		MS-LS4-3	1	0	2	3
		MS-LS4-4	2	1	0	3
		MS-LS4-5	1	0	0	1
		MS-LS4-6	1	0	1	2
Physical Sciences	PS1	MS-PS1-1	2	1	4	7
		MS-PS1-2	1	1	0	2
		MS-PS1-3	0	0	3	3
		MS-PS1-4	1	1	2	4
		MS-PS1-5	1	2	6	9
		MS-PS1-6	1	0	3	4
	PS2	MS-PS2-1	1	0	1	2
		MS-PS2-2	1	0	1	2
		MS-PS2-3	1	0	1	2
		MS-PS2-4	1	0	3	4
		MS-PS2-5	0	3	5	8
	PS3	MS-PS3-1	0	1	2	3
		MS-PS3-2	1	0	3	4
		MS-PS3-3	3	0	1	4
		MS-PS3-4	2	0	1	3
		MS-PS3-5	2	0	0	2
	PS4	MS-PS4-1	2	1	1	4
		MS-PS4-2	2	0	1	3
Total			**74**	**33**	**86**	**193**

[a]Other MOU states include Connecticut, Hawaii, Idaho, Oregon, Rhode Island, Utah, West Virginia, and Wyoming.

Appendix 2-N

Adaptive Algorithm Design

Table of Contents

Adaptive Item Selection Algorithm

1. Introduction, Background, and Definitions

This document describes the adaptive item selection algorithm. The item selection algorithm is designed to cover a standards-based blueprint, which may include content, cognitive complexity, and item type constraints. The item selection algorithm will also include:

- the ability to customize an item pool based on access constraints and screen items that have been previously viewed or may not be accessible for a given individual;
- a mechanism for inserting embedded field-test items; and
- a mechanism for delivering "segmented" tests in which separate parts of the test are administered in a fixed order.

This document describes the algorithm and the design for its implementation for the test delivery system (TDS). The implementation builds extensively on the algorithm implemented in the Cambium Assessment, Inc (CAI)'s TDS and incorporates substantial CAI intellectual property. CAI will release the algorithm and the implementation described here under the same open-source license under which the rest of the open-source system is released.

The general approach described here is based on a highly parameterized multiple-objective utility function. The objective function includes:

- a measure of content match to the blueprint;
- a measure of overall test information; and
- measures of test information for each reporting category on the test.

We define an objective function that measures an item's contribution to each of these objectives, weighting them to achieve the desired balance among them. Equation (1) sketches this objective function for a single item.

$$f_{ijt} = w_2 \frac{1}{\sum_{r=1}^{R} d_{rj}} \sum_{r=1}^{R} s_{rit} p_r d_{rj} + w_1 \sum_{k=1}^{K} q_k h_{1k}(v_{kijt}, V_{kit}, t_k) + w_0 h_0(u_{ijt}, U_{it}, t_0) \qquad (1)$$

where the term w represents user-supplied weights that assign relative importance to meeting each of the objectives d_{rj} indicates whether item j has the blueprint-specified feature r, and p_r is the user-supplied priority weight for feature r. The term s_{rit} is an adaptive control parameter that is described. In general, s_{rit} increases for features that have not met their designated minimum as the end of the test approaches.

The remainder of the terms represents an item's contribution to measurement precision:

- v_{kijt} is the value of item j toward reducing the measurement error for reporting category k for examinee i at selection t; and
- u_{ijt} is the value of item j in terms of reducing the overall measurement error for examinee i at selection t.

The terms U_{it} and V_{kit} represent the total information overall and on reporting category k, respectively.

The term q_k is a user-supplied priority weight associated with the precision of the score estimate for reporting category k. The terms t represent precision targets for the overall score (t_0) and each score reporting category score. The functions $h(.)$ are given by:

$$h_0(u_{ijt}, U_{it}, t_0) = \begin{cases} au_{ijt} \text{ if } U_{it} < t_0 \\ bu_{ijt} \text{ otherwise} \end{cases}$$

$$h_{1k}(v_{kijt}, V_{kit}, t_k) = \begin{cases} c_k v_{kijt} \text{ if } V_{kit} < t_k \\ d_k v_{kijt} \text{ otherwise} \end{cases}$$

Items can be selected to maximize the value of this function. This objective function can be manipulated to produce a pure, standards-free adaptive algorithm by setting w_2 to zero or a completely blueprint-driven test by setting $w_1 = w_0 = 0$. Adjusting the weights to optimize performance for a given item pool will enable users to maximize information subject to the constraint that the blueprint is virtually always met.

We note that the computations of the content values and information values generate values on very different scales, and that the scale of the content value varies as the test progresses. Therefore, we normalize both the information and content values before computing the value of Equation (1). This normalization is given by $x = \begin{cases} 1 \; if \; min = max \\ \frac{v-min}{max-min} \; otherwise \end{cases}$, where min and max represent the minimum and maximum, respectively, of the metric computed over the current set of items or item groups.

The remainder of this section describes the overall program flow, the form of the blueprint, and the various value calculations employed in the objective function. Subsequent sections describe the details of the selection algorithm.

1.1 BLUEPRINT

Each test will be described by a single blueprint for each segment of the test and will identify the order in which the segments appear. The blueprint will include:

- an indicator of whether the test is adaptive or fixed form;
- termination conditions for the segment, which are described in a subsequent section;
- a set of nested content constraints, each of which is expressed as:

  - the minimum number of items to be administered within the content category;
  - the maximum number of items to be administered within the content category;
  - an indication of whether the maximum should be deterministically enforced (a “strict” maximum);
  - a priority weight for the content category p_r;
  - an explicit indicator as to whether this content category is a reporting category; and
  - an explicit precision-priority weight (q_k) for each group identified as a reporting category.
- a set of non-nested content constraints, which are represented as:
  - a name for the collection of items meeting the constraint;
  - the minimum number of items to be administered from this group of items;
  - the maximum number of items to be administered from this group of items;
  - an indication of whether the maximum should be deterministically enforced (a “strict” maximum);
  - a priority weight for the group of items p_r;
  - an explicit indicator as to whether this named group will make up a reporting category; and
  - an explicit precision-priority weight (q_k) for each group identified as a reporting category.
  - The priority weights, p_r on the blueprint, can be used to express values in the blueprint match. Large weights on reporting categories paired with low (or zero) weights on the content categories below them may allow more flexibility to maximize information in a content category covering fewer fine-grained targets, while the reverse would mitigate toward more reliable coverage of finer-grained categories, with less content flexibility within reporting categories.

An example of a blueprint specification appears in Appendix J-1.

1.2 Content Value

Each item or item group will be characterized by its contribution to meeting the blueprint, given the items that have already been administered at any point. The contribution is based on the presence or absence of features specified in the blueprint and denoted by the term d in Equation (1). This section describes the computation of the content value.

1.2.1 Content Value for Single Items

For each constraint appearing in the blueprint (r), an item i either does or does not have the characteristic described by the constraint. For example, a constraint might require a minimum of four and a maximum of six algebra items. An item measuring algebra has the described characteristic, and an item measuring geometry, but algebra does not. To capture this constraint, we define the following:

- d_j is a feature vector in which the elements are d_{rj}, summarizing item j's contribution to meeting the blueprint. This feature vector includes content categories such as claims and targets as well as other features of the blueprint, such as Depth of Knowledge (DOK) and item type.
- S_{it} is a diagonal matrix, the diagonal elements of which are the adaptive control parameters s_{rit}.
- p is the vector containing the user-supplied priority weights p_r.

The scalar content value for an item is given by $C_{ijt} = d_j' S_{it} p$.

Letting z_{rit} represent the number of items with feature r administered to student i by iteration t, the value of the adaptive control parameters is:

$$s_{rit} = \begin{cases} m_{it}\left(2 - \dfrac{z_{rit}}{Min_r}\right) & if\ z_r < Min_r \\ 1 - \dfrac{z_{rit} - Min_r}{Max_r - Min_r} & if\ Min_r < z_{rit} < Max_r \\ (Max_r - z_{rit}) - 1 & if\ Max_r \leq z_{rit} \end{cases}$$

The blueprint defines the minimum (Min_r) and maximum (Max_r) number of items to be administered with each characteristic (r).

The term $m_{it} = \frac{T}{T-t}$ where T is the total test length. This has the effect of increasing the algorithm's preference for items that have not yet met their minimums as the end of the test nears and the opportunities to meet the minimum diminish.

This increases the likelihood of selecting items for content that has not met its minimum as the opportunities to do so are used up. The value s is highest for items with content that has not met its minimum, declines for items representing content for which the minimum number of items has been reached but the maximum has not, and turns negative for items representing content that has met the maximum.

1.2.2 Content Value for Sets of Items

Calculation of the content value of sets of items is complicated by two factors:

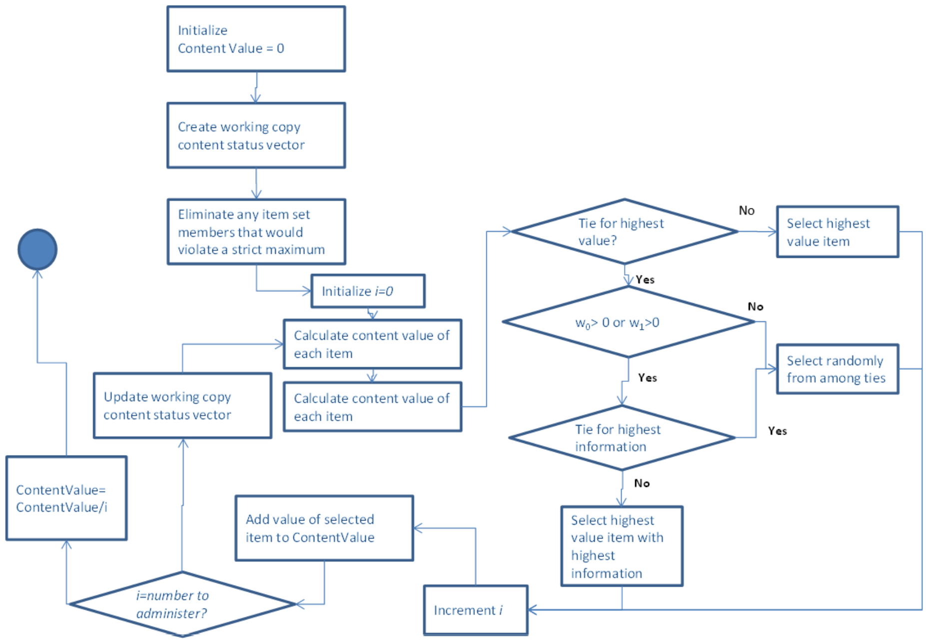
1. The desire to allow more items to be developed for each set and to have the most advantageous set of items administered.
2. The design objective of characterizing the information contribution of a set of items as the expected information over the working theta distribution for the examinee.

The former objective is believed to enhance the ability to satisfy highly constrained blueprints while still adapting to obtain good measurement for a broad range of students. The latter arises from the recognition that English Language Arts (ELA) tests will select one set of items at a time, without an opportunity to adapt once the passage has been selected.

The general approach involves successive selection of the highest content value item in the set until the indicated number of items in the set have been selected. Because the content value of an item changes with each selection, a temporary copy of the already-administered content vector for the examinee is updated with each selection such that subsequent selections reflect the items selected in previous iterations.

Exhibit A on the following page presents a flowchart for this calculation. Readers will note the check to determine whether $w_0 > 0$ or $w_1 > 0$. These weights, defined with Equation (1), identify the user-supplied importance of information optimization relative to blueprint optimization. In cases such as independent field tests, this weight may be set to zero, as it may not be desirable to make item administration dependent on the match to student performance. In more typical adaptive cases where item statistics will not be recalculated, favoring more informative items is generally better. The final measure of content value for the set of selected set of items is divided by the number of items selected to avoid a bias toward selection of sets with more items.

Exhibit A. Content Value Calculation for Item Sets



1.3 Information Value

Each item or item group also has value in terms of maximizing information, both overall and on reporting categories.

1.3.1 Individual Information Value

The information value associated with an item will be an approximation of information. The system will be designed to use generalized Item Response Theory (IRT) models; however, it will treat all items as though they offer equal measurement precision. This is the assumption made by the Rasch model, but in more general models, items known to offer better measurement are given preference by many algorithms. Subsequent algorithms are then required to control the exposure of the items that measure best. Ignoring the differences in slopes serves to eliminate this bias and help equalize exposure.

1.3.2 Binary Items

The approximate information value of a binary item will be characterized as $I_j(\theta) = p_j(\theta)(1 - p_j(\theta))$, where the slope parameters are artificially replaced with a constant.

1.3.3 Polytomous Items

In terms of information, the best polytomous item in the pool is the one that maximizes the expected information, $I_j(\theta)$. Formally, $I_j(\theta) > I_k(\theta)$ for all items $k \neq j$. The true value θ, however, remains unknown and is accessed only through an estimate, $\hat{\theta} \sim N(\bar{\theta}, \sigma_\theta)$. By definition of an expectation, the expected information $I_j(\theta) = \int I_j(t) f(t|\bar{\theta}, \sigma_\theta) dt$.

The intuition behind this result is illustrated in Exhibit B. In Exhibit B, each panel graphs the distribution of the estimate of θ for an examinee. The top panel assumes a polytomous item in which one step threshold (A1) matches the mean of the θ estimate distribution. In the bottom panel, neither step threshold matches the mean of the θ estimate distribution. The shaded area in each panel indicates the region in which the hypothetical item depicted in the panel provides more information. We see that approximately 2/3 of the probability density function is shaded in the lower panel, while the item depicted in the upper panel dominates in only about 1/3 of the cases. In this example, the item depicted in the lower panel has a much greater probability of maximizing the information from the item, despite the fact that the item in the upper panel has a threshold exactly matching the mean of the estimate distribution and the item in the lower panel does not.

Exhibit B. Two Example Items, with the Shaded Region Showing the Probability that the Item Maximizes Information for the Examinee Depicted

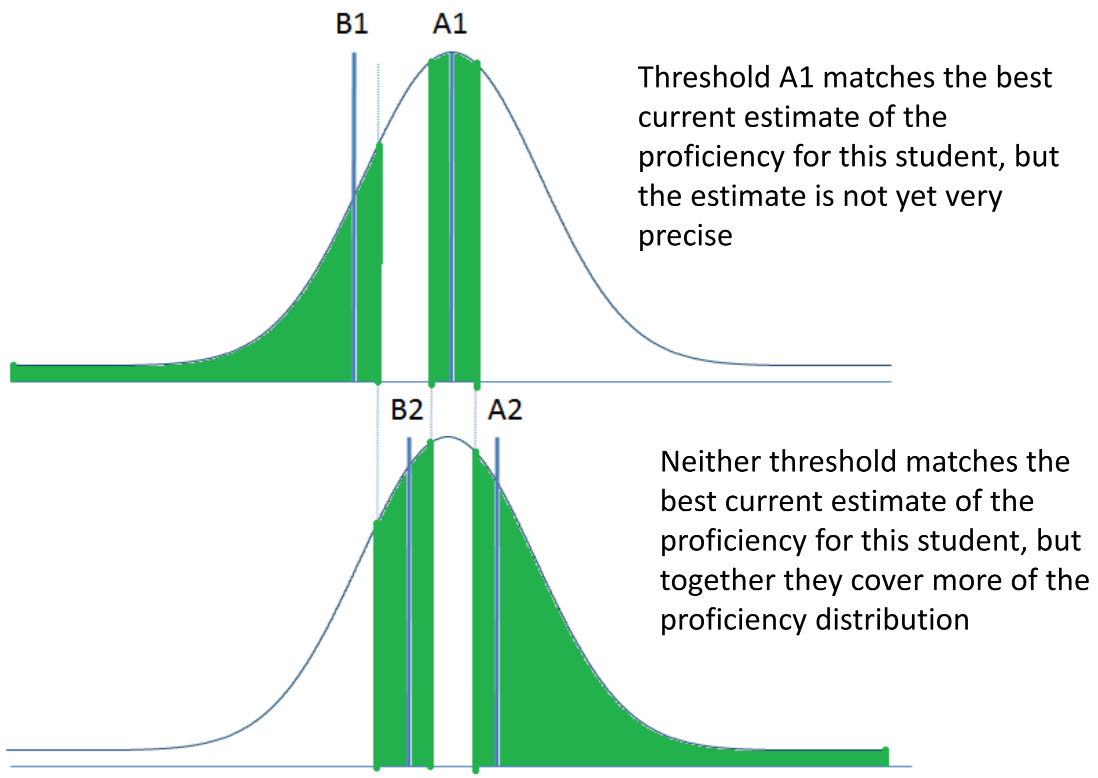
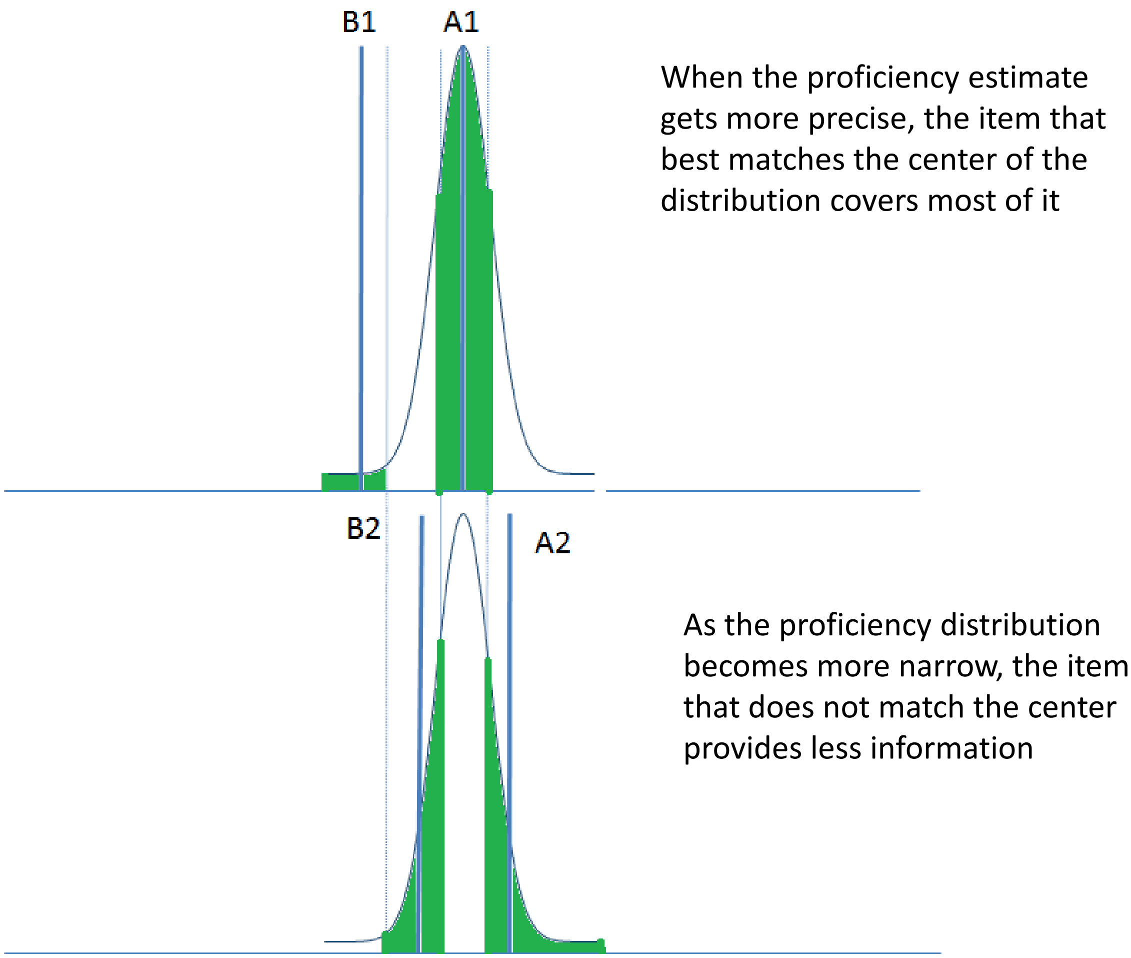


Exhibit C on the following page shows what happens to information as the estimate of this student's proficiency becomes more precise (later in the test). In this case, the item depicted in the top panel maximizes information about 65 to70 percent of the time, compared to about 30 to 35 percent for the item depicted in the lower panel. These are the same items depicted in the Exhibit B, but in this case, we are considering information for a student with a more precise current proficiency estimate.

Exhibit C. Two Example Items, with the Shaded Region Showing the Probability that the Item Maximizes Information for the Examinee Depicted



The approximate information value of polytomous items will be characterized as the expected information, specifically $E\left[I_j(\theta)|m_i, s_i\right] = \int \sum_{k=1}^{K} I_{jk}(t)\, p_j(k|t)\phi(t; m_i, s_i)dt$, where $I_{jk}(t)$ represents the information at t of response k to item j, $p_j(k|t)$ is the probability of response k to item j (artificially holding slope constant), given proficiency t, $\phi(.)$ represents the normal probability density function, and m_i and s_i represent the mean and standard deviation of examinee i's current estimated proficiency distribution.

We propose to use Gauss-Hermite quadrature with a small number of quadrature points (approximately five). Experiments show that we can complete this calculation for 1,000 items in fewer than 5 milliseconds, making it computationally reasonable.

As with the binary items, we propose to ignore the slope parameters to even exposure and avoid a bias toward the items with better measurement.

1.3.4 Item Group Information Value

Item groups differ from individual items in that a set of items will be selected for administration. Therefore, the goal is to maximize information across the working theta distribution. As with the polytomous items, we propose to use Gauss-Hermite quadrature to estimate the expected information of the item group.

In the case of multiple-item groups

$$E\left[I_g(\theta)|m_i, s_i\right] = \frac{1}{J_g}\int \sum_{j=1}^{J_g} I_{g(j)}(t)\, \phi(t; m_i, s_i)dt$$

Where $I_g(.)$ is the information from item group g, $I_{g(j)}$ is the information associated with item $j \in g$, for the J_g items in set g. In the case of polytomous items, we use the expected information, as described above.

2. Entry and Initialization

At startup, the system will

- create a custom item pool;
- initialize theta estimates for the overall score and each score point; and
- insert embedded field-test items.

2.1 Item Pool

At test startup, the system will generate a *custom item pool*, a string of item IDs for which the student is eligible. This item pool will include all items that

- are active in the system at test startup; and
- are not flagged as "access limited" for attributes associated with this student.

The list will be stored in ascending order of ID.

2.2 Adjust Segment Length

Custom item pools run the risk of being unable to meet segment blueprint minimums. To address this special case, the algorithm will adjust the blueprint to be consistent with the custom item pool. This capability becomes necessary when an accommodated item pool systematically excludes some content.

Let

$\boldsymbol{S}$ be the set of top-level content constraints in the hierarchical set of constraints, each consisting of the tuple (*name*, *min*, *max*, *n*);

$\boldsymbol{C}$ be the custom item pool, each element consisting of a set of content constraints $\boldsymbol{B}$;

$\boldsymbol{f}, \boldsymbol{p}$ integers represent item shortfall and pool count, respectively; and

$\boldsymbol{t}$ be the minimum required items on the segment.

For each s in $\boldsymbol{S}$, compute n as the sum of active operational items in $\boldsymbol{C}$ classified on the constraint.

$\boldsymbol{f}$ = summation over S $(min - n)$

$\boldsymbol{p}$ = summation over S (n)

if $t - f < p$, then $t = t - f$

2.3 Initialization of Starting Theta Estimates

The user will supply five pieces of information in the test configuration:

1. A default starting value if no other information is available

2. An indication whether prior scores on the same test should be used, if available
3. Optionally, the test ID of another test that can supply a starting value, along with
4. Slope and intercept parameters to adjust the scale of the value to transform it to the scale of the target test
5. A constant prior variance for use in calculation of working EAP scores

2.4 Insertion Of Embedded Field-Test Items

Each blueprint will specify

- the number of field-test items to be administered on each test;
- the first item position into which a field-test item may be inserted; and
- the last item position into which a field-test item may be inserted.

Upon startup, select randomly from among the field-test items or item sets until the system has selected the specified number of field-test items. If the items are in sets, the sets will be administered as a complete set, and this may lead to more than the specified number of items administered.

The probability of selection will be given by $p_j = \frac{\sum_{j=1}^{K} K_j}{\sum_{j=1}^{K} a_j K_j} a_j K_j \frac{m}{N_J}$, where

p_j represents the probability of selecting the item;

m is the targeted number of field-test items;

N_j is the total number of active items in the field-test pool;

K_j is the number of items in item set j; and

a_j is a user-supplied weight associated with each item (or item set) to adjust the relative probability of selection.

The a_j variables are included to allow for operational cases in which some items must complete field testing sooner or enter field testing later. While using this parameter presents some statistical risk, not doing so poses operational risks.

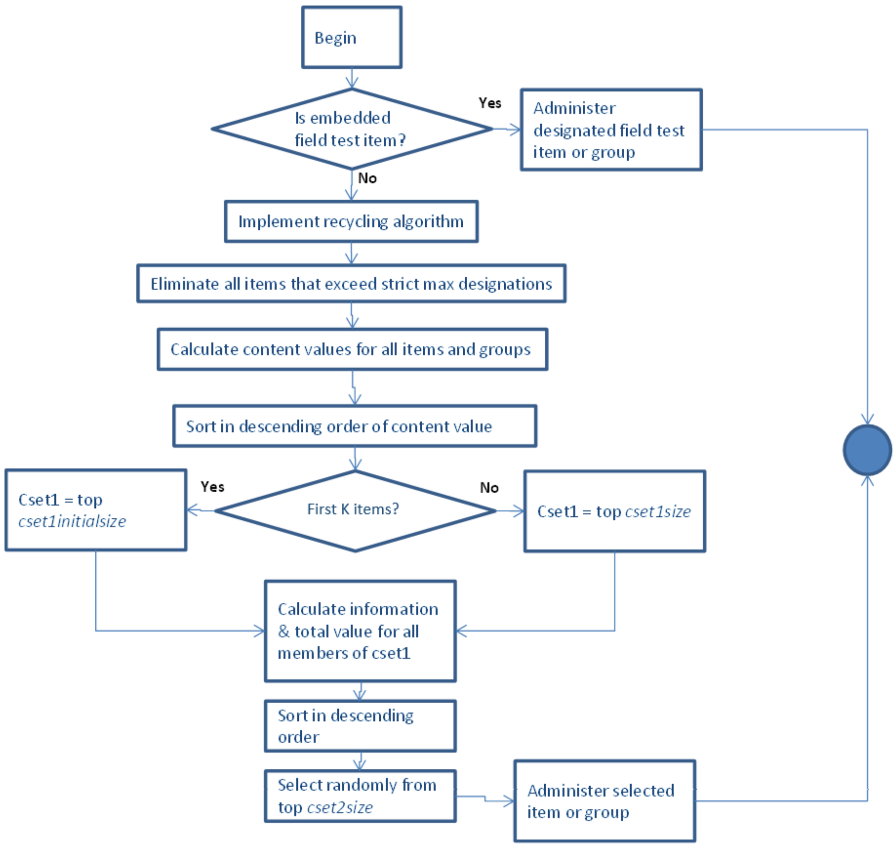
For each item set, generate a uniform random number r_j on the interval $\{0,1\}$. Sort the items in ascending order by $\frac{r_j}{p_j}$. Sequentially select items, summing the number of items in the set. Stop the selection of field-test items once $FTNMin \leq m \leq FTNMax = \sum_{j=0} K_j$.

Next, each item is assigned to a position on the test. To do so, select a starting position within $f - FTMax - FTMin$ positions from $FTMin$, where $FTMax$ is the maximum allowable position for field-test items and $FTMin$ is the minimum allowable position for field-test items. $FTNMin$ and $FTNMax$ refer to the minimum and maximum number of field-test items, respectively. Distribute the items evenly within these positions.

3. Item Selection

Exhibit D summarizes the item selection process. If the item position has been designated for a field-test item, administer that item. Otherwise, the adaptive algorithm kicks in.

Exhibit D. Summary of Item Selection Process



This approach is a "content first" approach designed to optimize match to blueprint. An alternative, "information first" approach, is possible. Under an information first approach, all items within a specified information range would be selected as the first set of candidates, and subsequent selection within that set would be based, in part, on content considerations. The engine is being designed so that future development could build such an algorithm using many of the calculations already available.

3.1 Trimming The Custom Item Pool

At each item selection, the active item pool is modified in four steps:

1. The custom item pool is intersected with the active item pool, resulting in a custom active item pool.
2. Items already administered on this test are removed from the custom active item pool.

3. Items that have been administered on prior tests are tentatively removed (see Section 3.2, Recycling Algorithm).
4. Items that measure content that has already exceeded a strict maximum are tentatively removed from the pool, removing entire sets containing items that meet this criterion.

3.2 RECYCLING ALGORITHM

When students are offered multiple opportunities to test, or when prior tests have been started and invalidated, students will have seen some of the items in the pool. The trimming of the item pool eliminates these items from the pool. It is possible that in such situations, the pool may no longer contain enough items to meet the blueprint.

Hence, items that have been seen on previous administrations may be returned to the pool. If there are not enough items remaining in the pool, the algorithm will recycle items (or item groups) with the required characteristic that is found in insufficient numbers. Working from the least recently administered group, items (or item groups) are reintroduced into the pool until the number of items with the required characteristics meets the minimum requirement. When item groups are recycled, the entire group is recycled rather than an individual item. Items administered on the current test are never recycled.

3.3 ADAPTIVE ITEM SELECTION

Selection of items will follow a common logic, whether the selection is for a single item or an item group. Item selection will proceed in the following three steps:

1. Select Candidate Set 1 (*cset1*).
    a. Calculate the content value of each item or item group.
    b. Sort the item groups in descending order of content value.
    c. Select the top *cset1size*, a user-supplied value that may vary by test.
2. Select Candidate Set 2 (*cset2*).
    a. Calculate the information values for each item group in *cset1*.
    b. Calculate the overall value of each item group in *cset1* as defined in Equation (1).
    c. Sort *cset2* in descending order of value.
    d. Select the top *cset2size* item groups, where *cset2size* is a user-supplied value that may vary by test.
3. Select the item or item group to be administered.
    a. Select randomly from *cset2* with uniform probability.

Note that a “pure adaptive” test, without regard to content constraints, can be achieved by setting *cset1size* to the size of the item pool and w_2, the weight associated with meeting content constraints in Equation (1), to zero. Similarly, linear on-the-fly tests can be constructed by setting w_0 and w_1 to zero.

3.4 Selection of The Initial Item

Selection of the initial item can affect item exposure. At the start of the test, all tests have no content already administered, so the items and item groups have the same content value for all examinees. In general, it is a good idea to spread the initial item selection over a wider range of content values. Therefore, we define an additional user-settable value, *cset1initialsize*, which is the size of Candidate Set 1 on the first K items only, where K is the number of reporting categories. Similarly, we define *cset2initialisize*.

3.5 Exposure Control

This algorithm uses randomization to control exposure and offers several parameters that can be adjusted to control the tradeoff between optimal item allocation and exposure control. The primary mechanism for controlling exposure is the random selection from $CSET2$, the set of items or item groups that best meet the content and information criteria. These represent the “top k” items, where k can be set. Larger values of k provide more exposure control at the expense of optional selection.

In addition to this mechanism, we avoid a bias toward items with higher measurement precision by treating all items as though they measured with equal precision by ignoring variation in the slope parameter. This has the effect of randomizing over items with differing slope parameters. Without this step, it would be necessary to have other *post hoc* explicit controls to avoid the overexposure of items with higher slope parameters, an approach that could lead to different test characteristics over the course of the testing window.

4. TERMINATION

The algorithm will have configurable termination conditions. These may include

- administering a minimum number of items in each reporting category and overall;
- achieving a target level of precision on the overall test score;
- achieving a target level of precision on all reporting categories; and
- achieving a score insufficiently distant from a specified score with sufficient precision (e.g., less than two standard errors below proficient). Cambium Assessment, Inc (CAI) envisions this being used in conjunction with other termination conditions to allow very high or very low achieving students to continue on to a segment that contains items from adjacent grades but barring other students from those segments.

We will define four user-defined flags indicating whether each of these is to be considered in the termination conditions (*TermCount*, *TermOverall*, *TermReporting*, *TermTooClose*). A fifth user-supplied value will indicate whether these are taken in conjunction or if satisfaction of any one of them will suffice (*TermAnd*). Reaching the minimum number of items is always a necessary condition for termination.

In addition, two conditions will each individually and independently cause termination of the test:

1. Administering the maximum number of items specified in the blueprint
2. Having no items in the pool left to administer

APPENDIX 1. DEFINITIONS OF USER-SETTABLE PARAMETERS

This appendix summarizes the user-settable parameters in the adaptive algorithm.

Parameter Name	Description	Entity Referred to by Subscript Index
w_0	Priority weight associated with overall information	N/A
w_1	Priority weight associated with reporting category information	N/A
w_2	Priority weight associated with match to blueprint	N/A
q_k	Priority weight associated with a specific reporting category	reporting categories
p_r	Priority weight associated with a feature specified in the blueprint (These inputs appear as a component of the blueprint.)	features specified in the blueprint
a	Parameter of the function $h(.)$ that controls the overall information weight when the information target has not yet been hit	N/A
b	Parameter of the function $h(.)$ that controls the overall information weight after the information target has been hit	N/A
c_k	Parameter of the function $h(.)$ that controls the information weight when the information target has not yet been hit for reporting category k	reporting categories
d_k	Parameter of the function $h(.)$ that controls the information weight after the information target has been hit for reporting category k	reporting categories
cset1size	Size of candidate pool based on contribution to blueprint match	N/A
cset1initialsize	Size of candidate pool based on contribution to blueprint match for the first K items or item sets selected	N/A
cset2size	Size of final candidate pool from which to select randomly	N/A
cset2initialsize	Size of candidate pool based on contribution to blueprint match and information for the first item or item set selected	
t_0	Target information for the overall test	N/A
t_k	Target information for reporting categories	reporting categories
startTheta	A default starting value if no other information is available	N/A
startPrevious	An indication of whether previous scores on the same test should be used, if available	N/A
startOther	The test ID of another test that can supply a starting value, along with startOtherSlope	N/A
startOtherSlope	Slope parameter to adjust the scale of the value to transform it to the scale of the target test	N/A

Parameter Name	Description	Entity Referred to by Subscript Index
startOtherInt	Intercept parameter to adjust the scale of the value to transform it to the scale of the target test	N/A
$FTMin$	Minimum position in which field-test items are allowed	N/A
$FTMax$	Maximum position in which field-test items are allowed	N/A
$FTNMin$	Target minimum number of field-test items	N/A
$FTNMax$	Target maximum number of field-test items	N/A
a_j	Weight adjustment for individual embedded field-test items used to increase or decrease their probability of selection	field-test items
AdaptiveCut	The overall score cutscore, usually proficiency, used in consideration of *TermTooClose*	
TooCloseSEs	The number of standard errors below which the difference is considered "too close" to the adaptive cut to proceed. In general, this will signal proceeding to a final segment that contains off-grade items.	
TermOverall	Flag indicating whether to use the overall information target as a termination criterion	N/A
TermReporting	Flag to indicate whether to use reporting category information target as a termination criterion	N/A
TermCount	Flag to indicate whether to use minimum test size as a termination condition	N/A
TermTooClose	Terminate if you are not sufficiently distant from the specified adaptive cut	
TermAnd	Flag to indicate whether the other termination conditions are to be taken separately or conjunctively	N/A

APPENDIX 2. SUPPORTING DATA STRUCTURES

Cambium Assessment, Inc (CAI) Cautions and Caveats

- Use of standard error termination conditions will likely cause inconsistencies between the blueprint content specifications, and the information criteria will cause unpredictable results, likely leading to failures to meet blueprint requirements.
- The field-test positioning algorithm outlined here is very simple and will lead to deterministic placement of field-test items.

ADDENDUM. ADJUSTMENTS TO THE USE OF ITEM CLUSTERS

Cambium Assessment, Inc (CAI) adjusted the adaptive algorithm to the use of item clusters as follows:

- Using marginal maximum likelihood estimator (MMLE) to update proficiency estimates, marginalizing out cluster effects.
- Normalizing the information by the number of assertions within an item, to avoid over-selection of item clusters and stand-alone items with more assertions.

Montana Science Assessment

2022–2023

Volume 3:
Setting Performance Standards

OFFICE OF PUBLIC INSTRUCTION
ELSIE ARNTZEN, STATE SUPERINTENDENT

TABLE OF CONTENTS

List of Tables

LIST OF FIGURES

LIST OF APPENDICES

1. Executive Summary

On September 16, 2016, the Montana Board of Public Education (BPE) adopted the new Montana Science Content Standards. The new standards employ a three-dimensional conceptualization of science understanding, including science and engineering practices, crosscutting concepts, and disciplinary core ideas. With the adoption of the Montana Science Content Standards, and the development of new statewide assessments to measure student achievement relative to those standards, the Montana Office of Public Instruction (OPI) convened a standard-setting workshop to recommend a system of performance standards for determining whether students have met the learning goals defined by the Montana Science Content Standards.

Under contract to OPI, Cambium Assessment, Inc. (CAI) conducted the standard-setting workshop to recommend performance standards for the Montana Science Assessment (MSA) in grades 5 and 8. The workshop was conducted remotely on August 2–3, 2022, following the performance-level descriptor (PLD) workshop that was conducted remotely on August 1, 2022.

The Montana Science Assessment (MSA) is designed to measure the attainment of the Montana Science Content Standards adopted by the Montana BPE. The assessment includes both item clusters and stand-alone items. Item clusters represent a series of interrelated student interactions directed toward describing, explaining, and predicting scientific phenomena. Stand-alone items are added to increase the test's coverage of the standards while limiting increases in testing time and burden on students and schools. Test items were developed by CAI, in conjunction with a group of states working to implement three-dimensional science standards. Test items were developed to ensure that each student is administered a test meeting all elements of the MSA blueprints, which were constructed to align with the Montana Science Content Standards.

Montana science educators, serving as standard-setting panelists, followed a rigorous standardized procedure to recommend performance standards demarcating each performance level. To recommend performance standards for the MSA, panelists participated in the Assertion-Mapping Procedure (AMP), an adaptation of the Item-Descriptor (ID) Matching procedure (Ferrara & Lewis, 2012). Consistent with ordered-item procedures generally (e.g., Mitzel, Lewis, Patz, & Green, 2001), workshop panelists reviewed and recommended performance standards using an ordered set of scoring assertions[1] derived from student interactions within items. Because the new science items—specifically the item clusters—represent multiple, interdependent interactions through which students engage in scientific phenomena, scoring assertions cannot be meaningfully evaluated independently of the item interactions from which they are derived. Thus, panelists were presented ordered scoring assertions for each item separately rather than for the test overall. Panelists mapped each scoring assertion to the most apt PLD.

Panelists reviewed PLDs describing the degree to which students have achieved the Montana Science Content Standards. Section 5.4.1 of this volume describes PLD development. OPI reviewed and revised Range PLDs before the standard-setting workshop. After reviewing the

[1] Scoring assertions articulate the evidence the student provides as a means to infer a specific skill or concept, which is aligned to content standards. In other words, scoring assertions capture each measurable action of an item and articulate what evidence the student has provided to infer a specific skill or concept.

Range PLDs, standard-setting panelists worked to identify the knowledge and skills characteristic of students just qualifying for entry into each performance level.

Working through the ordered scoring assertions for each item, panelists mapped each assertion into one of the four performance levels—Level 1 (Novice), Level 2 (Nearing Proficiency), Level 3 (Proficient), and Level 4 (Advanced). The mapping of scoring assertions was based on the consideration of test content. Panelists were provided additional contextual information, including the percentage of students who performed at or above the performance level associated with each assertion, as well as the projected National Assessment of Educational Progress (NAEP) science achievement levels of the assertion. The panelists performed the assertion mapping in two rounds of standard setting. Panelists' mapping of the scoring assertions was used to identify the location of the three performance standards used to classify student achievement—Level 2 (Nearing Proficiency), Level 3 (Proficient), and Level 4 (Advanced). Following Round 1, panelists were provided with feedback about the mappings of their fellow panelists and discussed their mappings as a group. The performance standards from Round 2 showed good convergence within each grade-level panel, indicating panelists had a good understanding of the PLDs and the AMP. In both grades, for a specific performance level, the performance standards resulted in reasonable percentages. As the standard-setting results from Round 2 concluded the recommended performance standards, no modifications to the performance standards were needed.

Thirty Montana science educators[2] were selected to serve as science standard-setting panelists, with 14 participants for the grade 5 panel and 16 participants for the grade 8 panel. The panelists represented a group of experienced teachers and curriculum specialists, as well as district administrators and other stakeholders. The composition of the panel ensured that a diverse range of perspectives and deep experience with the Montana Science Content Standards contributed to the standard-setting process.

1.1 STANDARD-SETTING WORKSHOP

1.1.1 Overall Structure of the Workshop

The key features of the workshops included the following:

- The standard-setting procedure produced three recommended performance standards (Level 2 – Nearing Proficiency, Level 3 – Proficient, and Level 4 – Advanced) that will be used to classify student performance on the MSA in grades 5 and 8.
- Panelists recommended performance standards in two rounds.
- Contextual information, including the percentage of students who performed at or above the performance level value associated with each individual assertion (impact data) and the projected NAEP science performance levels of each assertion (benchmark information), was provided to panelists as part of their review of the ordered assertions.
- The standard-setting workshop was conducted using the CAI's online standard-setting tool. Because the workshop was conducted remotely, each panelist accessed the tool using their own devices. At the end of the workshop, panelists completed online workshop evaluations

[2] See Section 5.3.4, Educator Participants for more information on the panelists.

independently, in which they described and evaluated their experience taking part in the standard setting.

1.1.2 Results of the Standard-Setting Workshop

Table 1 displays the performance standards recommended by the standard-setting panelists.

Table 1. Performance Standards Recommended for Science

Grade	Level 2 Nearing Proficiency	Level 3 Proficient	Level 4 Advanced
5	477	506	531
8	781	808	826

Table 2 indicates the percentage of students that will reach or exceed each performance standard in 2022. Figure 1 represents those values graphically.

Table 2. Percentage of Students Reaching or Exceeding Each Recommended Science Performance Standard in 2022

Grade	Level 2 Nearing Proficiency	Level 3 Proficient	Level 4 Advanced
5	80	43	13
8	76	40	15

Figure 1. Percentage of Students Reaching or Exceeding Each Recommended Science Performance Standard in 2022

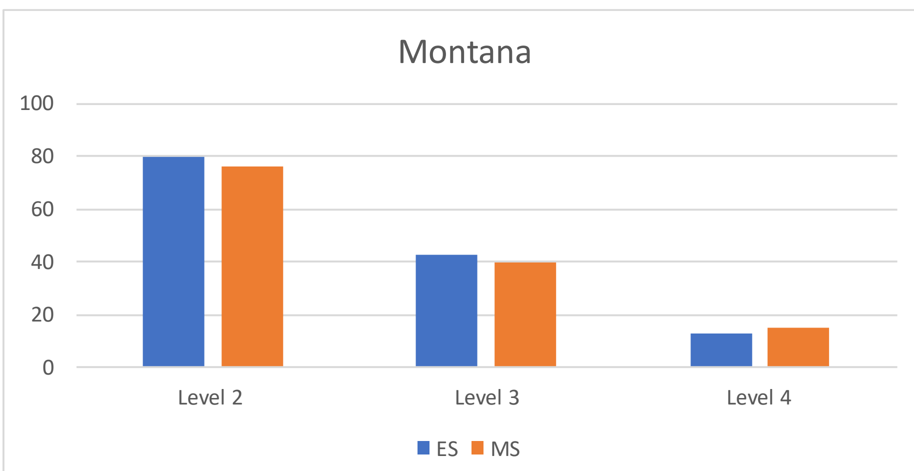
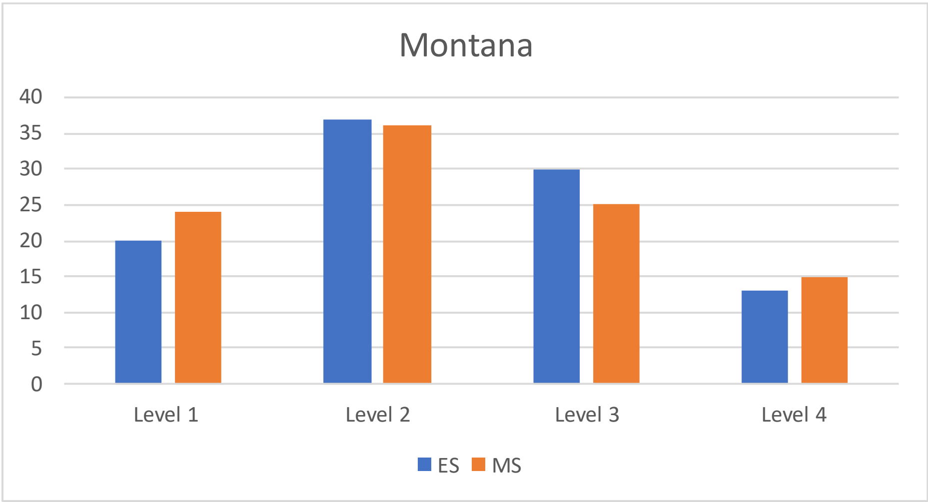


Table 3 indicates the percentage of students classified within each of the performance levels in 2022. The values are displayed graphically in Figure 2.

Table 3. Percentage of Students Classified Within Each Science Performance Level in 2022

Grade	Level 1 Novice	Level 2 Nearing Proficiency	Level 3 Proficient	Level 4 Advanced
5	20	37	30	13
8	24	36	25	15

Figure 2. Percentage of Students Classified Within Each Science Performance Level in 2022



2. Introduction

Montana adopted a set of three-dimensional science standards as the new Montana Science Content Standards in September 2016. The Montana Office of Public Instruction (OPI) and its assessment vendor, Cambium Assessment, Inc. (CAI), developed and administered a new assessment to measure the new standards. Field tested in 2020–2021 and administered operationally for the first time in 2021–2022, the Montana Science Assessment (MSA) measures the science knowledge and skills of Montana students in grades 5 and 8.

Montana provides information about the science assessments at: https://opi.mt.gov/Leadership/Assessment-Accountability/MontCAS/Required-Assessments/Montana-Science-Assessment-FAQ.

New tests require new performance standards to link performance on the test to the content standards. OPI contracted with CAI to establish cut scores for the new tests. To fulfill this responsibility, CAI implemented a technically sound methodology; provided training on standard setting to all participants; oversaw the process; computed real-time feedback data to inform the process; and produced a technical report documenting the method, approach, process, and outcomes. Performance standards were recommended for grades 5 and grade 8 in August 2022.

The purpose of this report volume is to detail the standard-setting process for the MSA and resulting performance standard recommendations.

3. THE MONTANA SCIENCE CONTENT STANDARDS

The Montana Science Assessment (MSA) assesses the learning objectives described by the Montana Science Content Standards, adopted by Montana in 2016.

The three-dimensional science standards, which are based on *A Framework for K–12 Science Education* (National Research Council, 2012), reflect the latest research and advances in modern science education and differ from previous science standards in multiple ways. First, rather than describe general knowledge and skills that students should know and be able to do, they describe specific performances that demonstrate what students know and can do. The Montana Science Content Standards refer to these performed knowledge and skills as *performance expectations* (PEs). Second, the Montana Science Content Standards are intentionally multi-dimensional. Each PE incorporates all three dimensions from *A Framework for K–12 Science Education* (National Research Council, 2012): a science or engineering practice, a disciplinary core idea, and a crosscutting concept. Another unique feature of the Montana Science Content Standards is the assumption that students should learn all science disciplines, rather than select a few, as is traditionally done in some schools.

Figure 3 shows the structure of the Montana Science Content Standards for a single grade 5 PE, 5-PS1-1.

Figure 3. Structure of the Montana Science Content Standards Performance Expectations

Students who demonstrate understanding can:

5-PS1-1. **Develop a model to describe that matter is made of particles too small to be seen.** [Clarification Statement: Examples of evidence supporting a model could include adding air to expand a basketball, compressing air in a syringe, dissolving sugar in water, and evaporating salt water.] [*Assessment Boundary: Assessment does not include the atomic-scale mechanism of evaporation and condensation or defining the unseen particles.*]

The performance expectation above was developed using the following elements from the NRC document *A Framework for K-12 Science Education*:

Science and Engineering Practices	Disciplinary Core Ideas	Crosscutting Concepts
Developing and Using Models Modeling in 3–5 builds on K–2 experiences and progresses to building and revising simple models and using models to represent events and design solutions. • Use models to describe phenomena.	**PS1.A: Structure and Properties of Matter** • Matter of any type can be subdivided into particles that are too small to see, but even then the matter still exists and can be detected by other means. A model showing that gases are made from matter particles that are too small to see and are moving freely around in space can explain many observations, including the inflation and shape of a balloon and the effects of air on larger particles or objects.	**Scale, Proportion, and Quantity** • Natural objects exist from the very small to the immensely large.

Connections to other DCIs in fifth grade: N/A

Articulation of DCIs across grade-levels:
2.PS1.A ; **MS.PS1.A**

Common Core State Standards Connections:

ELA/Literacy -

RI.5.7	Draw on information from multiple print or digital sources, demonstrating the ability to locate an answer to a question quickly or to solve a problem efficiently. *(5-PS1-1)*

Mathematics -

MP.2	Reason abstractly and quantitatively. *(5-PS1-1)*
MP.4	Model with mathematics. *(5-PS1-1)*
5.NBT.A.1	Explain patterns in the number of zeros of the product when multiplying a number by powers of 10, and explain patterns in the placement of the decimal point when a decimal is multiplied or divided by a power of 10. Use whole-number exponents to denote powers of 10. *(5-PS1-1)*
5.NF.B.7	Apply and extend previous understandings of division to divide unit fractions by whole numbers and whole numbers by unit fractions. *(5-PS1-1)*
5.MD.C.3	Recognize volume as an attribute of solid figures and understand concepts of volume measurement. *(5-PS1-1)*
5.MD.C.4	Measure volumes by counting unit cubes, using cubic cm, cubic in, cubic ft, and improvised units. *(5-PS1-1)*

* The performance expectations marked with an asterisk integrate traditional science content with engineering through a Practice or Disciplinary Core Idea.

Source. https://www.nextgenscience.org/pe/5-ps1-1-matter-and-its-interactions.

Information about Montana Science Content Standards is available at: https://opi.mt.gov/Educators/Teaching-Learning/K-12-Content-Standards/Science-Standards#88989723-science-content-standards.

4. Montana Science Assessment

Due to the unique features of the three-dimensional Montana Science Content Standards, items and tests based on these standards, such as the Montana Science Assessment (MSA, must also incorporate similarly unique features. The most impactful of these changes is that new science tests are multi-dimensional and are thus made up mostly of *item clusters* representing a series of interrelated student interactions directed toward describing, explaining, and predicting scientific phenomena.

4.1 Item Clusters and Stand-Alone Items

There are two types of items: item clusters and stand-alone items. An item cluster includes a phenomenon-based stimulus and a series of interactions that allow the student to demonstrate their mastery of the performance expectation (PE) by explaining the phenomenon or designing a solution to a presented engineering problem. The expectation is that item clusters will take students approximately 10 to 12 minutes to complete. Each stimulus ends with a task statement that provides the goal or understanding the student should reach. For example, "In the questions that follow, you will analyze what happens to the train when the brakes are applied." The student may explain, model, investigate, and/or create designs using the knowledge, skills, and abilities described by the PE. For example, in Figure 3, proficiency in this single PE requires activities that demonstrate the ability to analyze and evaluate data, the knowledge of properties and purposes of different forms of matter, and the application of experimental cause and effect. All interactions within an item cluster address the phenomenon presented in the stimulus. Item clusters contain between four and eight interactions.

Most states also utilize stand-alone items. Stand-alone items increase the number of covered PEs per student while being much quicker to complete than item clusters. Incorporating stand-alone items allows the blueprint to cover a greater number of PEs within a limited time. Stand-alone items are also phenomenon-based, contain only one or two interactions, and take students one to three minutes to complete in general.

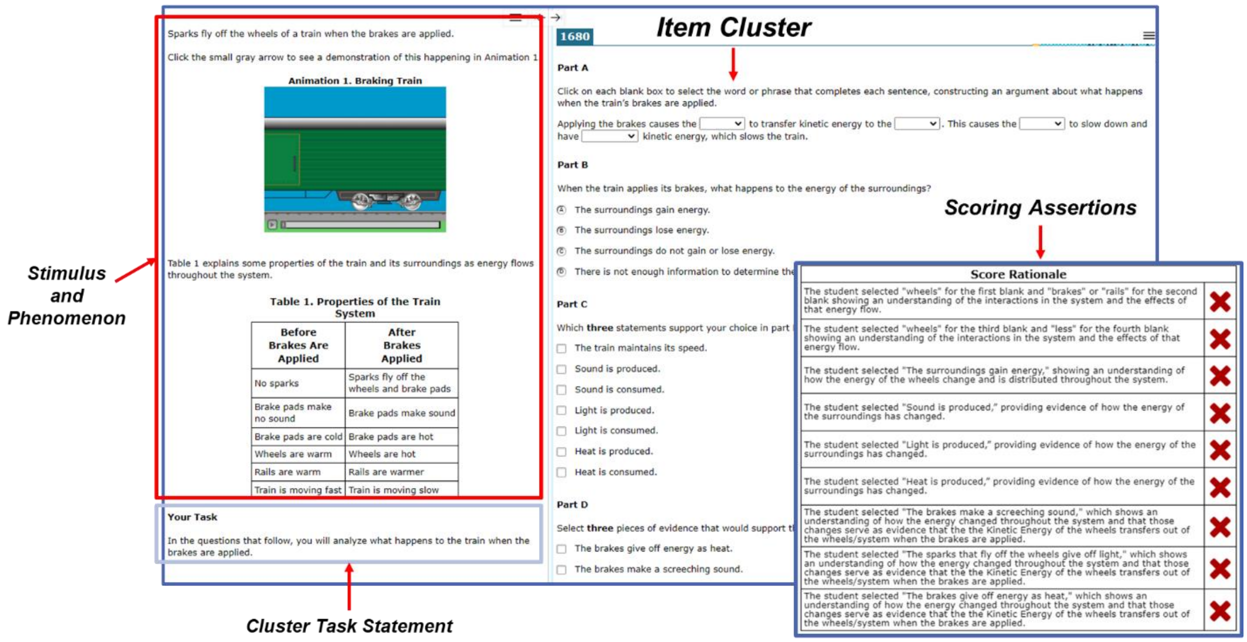
Both item types may use any of the available interaction types, including selected response, multi-select, table match, external copy, edit in-line choice, grids, and/or simulations of scientific investigations. For additional information on interaction types, refer to Volume 2, Appendix C, Style Guide for Science Items, of this technical report.

4.2 Scoring Assertions

Each item cluster and stand-alone item assumes a series of explicit assertions about the knowledge and skills that a student demonstrates based on specific features of the student's responses across multiple interactions. *Scoring assertions* capture each measurable action and articulate what evidence the student has provided to infer a specific skill or concept. Some stand-alone items have more than one scoring assertion, while all item clusters have multiple scoring assertions.

Figure 4 illustrates an item cluster and associated scoring assertions.

Figure 4. Example of the Three-Dimensional Science Item Cluster and Scoring Assertions

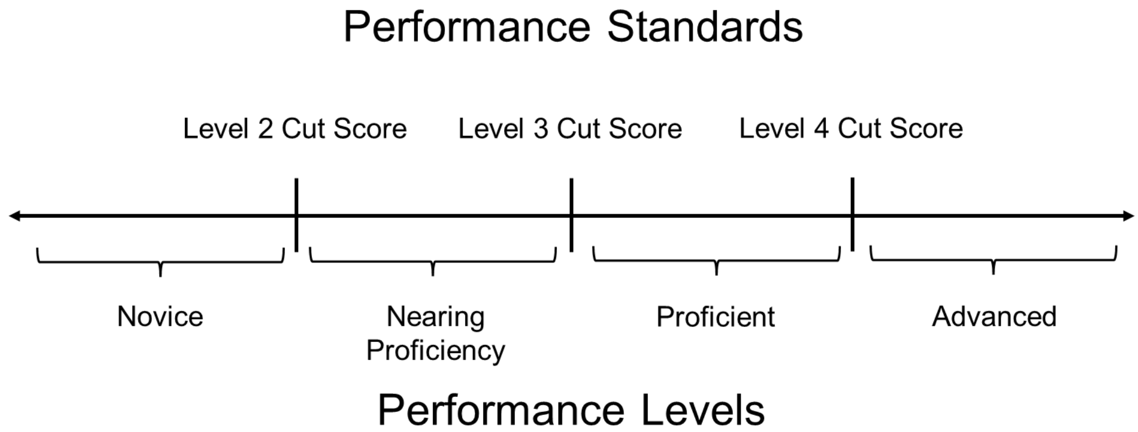


5. STANDARD SETTING

Thirty educators from Montana convened remotely on August 2–3, 2022, to complete two rounds of standard setting to recommend three performance standards for the Montana Science Assessment (MSA).

Standard setting is the process used to define performance on the test. Performance levels are defined by performance standards, or *cut scores*, that specify how much of the performance expectations (PEs) students must know and be able to do in order to meet the minimum for each performance level. As shown in Figure 5, three performance standards are sufficient to define Montana’s four performance levels.

Figure 5. Three Performance Standards Defining Montana's Four Performance Levels



The cut scores are derived from the knowledge and skills measured by the test item-scoring assertions that students at each performance level are expected to be able to receive credit.

5.1 The Assertion-Mapping Procedure

A modification of traditional approaches to setting performance standards is necessary for tests based on the Montana Science Content Standards due to the structure of the PEs, and subsequently, the structure of test items assessing the PEs. While traditional tests and measurement models assume unidimensionality, tests based on the Montana Science Content Standards adopt a three-dimensional conceptualization of science understanding. Each item cluster or stand-alone item aligns with a science practice, one or more crosscutting concepts, and one disciplinary core idea. Accordingly, the new science assessments comprise mostly item clusters representing a series of interrelated student interactions directed toward describing, explaining, and predicting scientific phenomena. Some stand-alone items are added to increase the test's coverage of the standards without also increasing testing time or testing burden.

Within each item, a series of explicit assertions are made regarding the knowledge and skills that a student has demonstrated based on specific features of the student's responses across multiple interactions. For example, students may correctly graph data points indicating that they can construct a graph showing the relationship between two variables but may make an incorrect inference regarding the relationship between the two variables, thereby not supporting the assertion that they can interpret relationships expressed graphically.

While some other assessments, especially English language arts (ELA), comprise items probing a common stimulus, the degree of interdependence among such items is limited and student performance on such items can be evaluated independently of student performance on other items within the stimulus set. This is not the case with the new science items, which may, for example, involve multiple steps in which students interact with products of previous steps. However, unlike traditional stimulus- or passage-based items, the conditional dependencies between the interactions and resulting assertions of an item cluster are too substantial to ignore because those item interactions and assertions are more intrinsically related to each other. The interdependence of student interactions within items has consequences both for scoring and recommending performance standards.

To account for the cluster-specific variation of related item clusters, additional dimensions can be added to the item response theory (IRT) model. Typically, these are nuisance dimensions unrelated to student ability. Examples of IRT models that follow this approach are the bi-factor model (Gibbons & Hedeker, 1992) and the testlet model (Bradlow, Wainer, & Wang, 1999). The testlet model is a special case of the bi-factor model (Rijmen, 2010).

Because the item clusters represent performance tasks, the Body of Work (BoW) method (Kingston, Kahl, Sweeny, & Bay, 2001) could also be appropriate for recommending performance standards. However, the BoW method is manageable only with small numbers of performance tasks and quickly becomes onerous when the number of item clusters approaches 10 or more.

Skaggs, Hein, & Awuor (2007) proposed a standard setting method called the Single-Passage Bookmark method to address challenges presented by passage-based assessments. This method is a variation of the traditional Bookmark method (e.g., Mitzel, Lewis, Patz, & Green, 2001) in which individual ordered item booklets (OIBs) are created for each set of items associated with a passage. Items within each OIB are arranged in order of difficulty. The task of the panelists is to place a bookmark in each OIB as opposed to a single OIB in the traditional Bookmark method. Even though this method showed promise, one limitation and concern expressed by the authors is whether this method can be applied to derive two or more standards. To address these challenges, Cambium Assessment, Inc. (CAI) psychometricians designed a new method for setting performance standards on cluster-based assessments. CAI implemented this method for the New Hampshire, Utah, and West Virginia statewide assessments in 2018; for the Connecticut, Oregon, and the joint Multi-State Science Assessment (MSSA) for Rhode Island and Vermont in 2019; and for the Hawaii, South Dakota, and Utah statewide assessments in 2021. In 2022, the method was also implemented for the Idaho, U.S. Virgin Islands, and Wyoming statewide assessments.

The test-centered Assertion-Mapping Procedure (AMP) is an adaptation of the Item-Descriptor (ID) Matching procedure (Ferrara & Lewis, 2012) that preserves the integrity of the item clusters while also taking advantage of ordered-item procedures such as the Bookmark method used frequently for other accountability tests (Rijmen, Cohen, Butcher, & Farley, 2018).

The main distinction between AMP and the Single-Passage Bookmark method is that the panelists evaluate scoring assertions rather than individual items. Scoring assertions are not test items, but inferences that are supported (or not supported) by students' responses in one or more interactions within an item cluster or stand-alone item. Because item clusters represent multiple, interdependent interactions through which students engage in scientific phenomena, scoring assertions cannot be meaningfully evaluated independently of the item from which they are derived. Therefore, the scoring assertions from the same item cluster or stand-alone item are always presented together. Within each item cluster or stand-alone item, scoring assertions are ordered by difficulty (i.e., the IRT difficulty parameter) consistent with the Single-Passage Bookmark method. One can think of the resulting booklet as consisting of different chapters, where each chapter represents an item cluster or stand-alone item. Within each chapter, the (ordered) pages represent scoring assertions. As in ID Matching, panelists are asked to map each scoring assertion to the most apt performance-level descriptor during two rounds of standard setting. As with the Bookmark method, assertion mappings are made independently with the goal of convergence over two rounds of rating, rather than consensus.[3]

5.2 Workshop Structure

One large virtual meeting room served as an all-participant training room. This room broke into two separate virtual working rooms, one for each set of test-level panels, after the all-group orientation. As shown in Figure 6, two separate panels set performance standards for each test.

Figure 6. Virtual Workshop Panels, per Room

Table 4. Table summarizes the composition of the tables and the number of facilitators and panelists assigned to each. The 30 standard-setting participants included table leaders and panelists from Montana who taught in the content area and grade for which standards were being set.

Table 4. Table Assignments

Room	Test	Tables and Table Leaders (One per Table)	Panelists (per Table)	Facilitators
1	Grade 5	2	6 / 8	Hibbah Haddam Olivia Francois
2	Grade 8	2	9 / 7	Matthew Davis Jared Taylor

[3] CAI historically implements two rounds of standard setting as best practice in the Bookmark method and extends this practice to the AMP. In addition to lessening the panelists' burden of needing to repeat a cognitively demanding task for a third time, using two rounds introduces significant cost efficiency by reducing the number of days needed for standard setting. Panels typically converge in Round 2, and panelists completing two rounds report levels of confidence in the outcomes that are similar to the confidence expressed by panelists participating in three rounds. Psychometric evaluation of the reliability and variability in results from two and three rounds are generally consistent. CAI has used two rounds in standard setting in more than 17 states and 38 assessments, beginning in 2001 with the enactment of the No Child Left Behind (NCLB) Act.

5.3 Participants and Roles

5.3.1 Montana Office of Public Instruction Staff

Staff from the Montana Office of Public Instruction (OPI) were present throughout the process and provided overall policy context and answered any policy questions that arose.

From OPI, attendees included:

- Ashley McGrath, Assessment Director
- Katie Murnion, Assessment Specialist
- Chris Noel, Teaching and Learning Senior Manager
- Michelle McCarthy, Science Instructional Coordinator
- Jennifer Stadum, Indian Education Specialist

5.3.2 Cambium Assessment, Inc. Staff

CAI facilitated the workshop and each of the content-area rooms, provided psychometric and statistical support, and oversaw technical setup and logistics. CAI team members were highly qualified to lead the workshop and conduct analyses, and included the following:

- Dr. Stephan Ahadi, Managing Director of Psychometrics, facilitated and oversaw all AMP processes and tasks and provided training to participants.
- Dr. Zhongtian Lin, Senior Psychometrician, provided psychometric analyses.
- Alesha Ballman, Psychometric Project Coordinator, oversaw analytics technology and psychometrics.
- Kylie Dennis and Sydney Brabble, Psychometric Support Assistants, provided support as needed.
- Emilia Birdsall and Evelyn Chester, Program Management Team, managed process and logistics throughout the meeting.
- Floyd Helm, Brandon Osborne, Mark Palomo, System Support Agents, troubleshot technology during the workshop.

5.3.3 Room Facilitators

Two CAI facilitators guided the process in each test-level room. Facilitators were content experts (e.g., hold degrees in science and/or science education and have multiple years' experience in writing science items and content development), experienced in leading standard-setting processes, had led standard-setting processes before, and could answer any questions about the workshop or about the items or what the items were intended to measure. They also monitored time and motivated panelists to complete tasks within the scheduled time. Facilitators included the following:

- Hibbah Haddam and Olivia Francois facilitated the science grade 5 panel
- Matthew Davis and Jared Taylor facilitated the science grade 8 panel

Each facilitator was trained to be extensively knowledgeable of the constructs, processes, and technologies used in standard setting.

Educator Participants

To establish performance standards, OPI recruited a set of participants from across the state. Panelists included science teachers, administrators, and representatives from other stakeholder
5.3.4 groups (e.g., parents, college faculty) to ensure that a range of perspectives contributed to the standard-setting process and product. In recruiting panelists, OPI targeted the recruitment of participants to be representative of the gender and geographic representation of Montana's teacher population. All participants also had to be familiar with the science content and test.

OPI selected classroom teachers from the resulting potential panelist pool and invited them to participate in the workshop. Overall, the standard-setting workshop panelists were predominantly female (93%) and White (97%). Represented stakeholder groups included Coaches, Special Education Teachers, Specialists, Parents, and General Education Teachers, with General Education Teachers comprising 62% of the panels overall. Most panelists taught in the grades to which they were assigned to set standards. School district areas included rural (50%), suburban (17%), and urban (27%), and were small (30%), medium (33%), and large (30%). Table 5 summarizes the characteristics of the panels.

Table 5. Panelist Characteristics

	Percentage of Panelists by Panel		
	Science Grade 5	*Science Grade 8*	*Overall*
Characteristics			
Male, Female	0%, 100%	13%, 87%	7%, 93%
White, Non-White	100%, 0%	94%, 6%	97%, 3%
Stakeholder Group			
General Education Teacher	63%	62%	62%
Specialist	0%	10%	5%
Coach	0%	5%	3%
Administrator	0%	10%	5%
Special Education Teacher	13%	0%	5%
English Learner Teacher	0%	0%	0%
Higher Education	0%	0%	0%
Parent	6%	5%	5%
Other[a]	19%	10%	14%
Current Position			
School	71%	50%	60%

	Percentage of Panelists by Panel		
	Science Grade 5	*Science Grade 8*	*Overall*
School, District	7%	44%	27%
Other[b]	21%	6%	13%
District Size			
Large	21%	38%	30%
Medium	50%	19%	33%
Small	14%	44%	30%
Not Applicable	14%	0%	7%
District Urbanicity			
Urban	29%	25%	27%
Suburban	7%	25%	17%
Rural	50%	50%	50%
Not Applicable	14%	0%	7%
Primary Grades Taught			
Preschool	7%	19%	13%
Kindergarten	14%	25%	20%
1st Grade	14%	19%	17%
2nd Grade	14%	19%	17%
3rd Grade	21%	19%	20%
4th Grade	36%	19%	27%
5th Grade	43%	25%	33%
6th Grade	21%	31%	27%
7th Grade	14%	38%	27%
8th Grade	36%	44%	40%
9th Grade	7%	31%	20%
10th Grade	7%	31%	20%
11th Grade	7%	25%	17%
12th Grade	7%	25%	17%
College	0%	0%	0%
Not Applicable	14%	0%	7%

Note. [a]Other Stakeholder Groups include Co-Teacher, Mentor, Naturalist Teacher, Retired, and Supervising Teacher. [b]Other Current Position Location includes the Montana Natural History Center, Missoula, MT, and the Montana Audubon Center, Billings, MT.

For the results of any judgment-based method to be valid, the judgments must be made by individuals who are qualified to make them. Participants in the MSA standard-setting workshop were highly qualified. They brought a variety of experience and expertise. Overall, 70% of panelists had earned a master's degree. More than half (77%) had taught in their assigned panel's grade and subject for one to 10 years, while 7% had taught it for more than 20 years. The average time teaching the Montana Science Content Standards was nearly seven years. Many had experience teaching special populations; 34% taught students eligible to receive free or reduced-price lunch, 27% taught English learners (ELs), and 39% taught students on an Individual Education Plan (IEP). Table 6 summarizes the qualifications of the panels.

Table 6. Panelist Qualifications

	Percentage of Panelists by Panel		
	Science Grade 5	*Science Grade 8*	*Overall*
Highest Degree			
Bachelor	29%	31%	30%
Master	71%	69%	70%
Doctoral	0%	0%	0%
Years Teaching Experience			
None	7%	0%	3%
Less than 1 year	0%	0%	0%
1–5 years	7%	31%	20%
6–10 years	14%	19%	17%
11–15 years	50%	19%	33%
16–20 years	0%	13%	7%
More than 20 years	21%	19%	20%
Years Teaching Experience in Assigned Grade			
None	29%	13%	20%
Less than 1 year	0%	0%	0%
1–5 years	36%	50%	43%
6–10 years	21%	6%	13%
11–15 years	7%	13%	10%
16–20 years	0%	13%	7%
More than 20 years	7%	6%	7%
Subject Areas Currently Teaching[a]			
ELA	19%	13%	16%
Mathematics	22%	19%	21%
Social Studies	14%	13%	13%
Science	33%	48%	40%
Not Applicable	6%	0%	3%
Other[b]	6%	6%	6%
Other Professional Experience in Education	50%	56%	53%
Years Professional Experience in Education			
None	50%	44%	47%
Less than 1 year	0%	0%	0%
1–5 years	29%	38%	33%
6–10 years	14%	13%	13%
11–15 years	0%	0%	0%
16–20 years	0%	6%	3%
More than 20 years	7%	0%	3%
Experience Teaching Special Student Populations			

	Percentage of Panelists by Panel		
	Science Grade 5	*Science Grade 8*	*Overall*
Students eligible to receive free/reduced price lunch	35%	33%	34%
ELs	29%	26%	27%
Students on an IEP	35%	41%	39%
Average Years Teaching the Montana Science Content Standards	9	5	7

Note. [a]The total sums to more than 100% for "Subject Areas Currently Teaching" as many participants taught multiple subjects. [b]Other Subject Areas Currently Teaching includes Art, Health, Social Emotional Learning (SEL), and Teacher Preparation Strategies Across Content (mathematics, science, and social studies).

Appendix 3-A, Standard-Setting Panelist Characteristics, provides additional information about the individuals participating in the standard-setting workshop.

5.3.5 Table Leaders

Volunteers from the participant pool served as table leaders, who were intimately familiar with students and the subject matter. The day prior to the standard-setting workshop, the group of Montana educators selected to be the table leaders convened to review, revise, and approve the range Performance-Level Descriptors (PLDs) (see Appendix C). During the standard-setting workshop, table leaders served as panelists. In addition, table leaders had the responsibility of participating in the moderation session (see Section 5.8.4, Moderation).

5.4 MATERIALS

5.4.1 Performance-Level Descriptors

With the adoption of the new standards in science, and the development of new statewide assessments to measure performance of those standards, OPI must adopt a similar system of performance, or performance standards, to determine whether students have met the learning goals defined by the new standards in science.

Determining the nature of the categories into which students are classified is a prerequisite to standard setting. These categories, or performance levels, are associated with PLDs that define the content-area knowledge, skills, and processes that students at each performance level can demonstrate.

PLDs link the content standards (PEs) to the performance standards. There are four types of PLDs (Egan, Schneider, & Ferrara, 2012):

1. **Policy PLDs.** These are brief descriptions of each performance level that do not vary across grade or content area.
2. **Range PLDs.** Provided to panelists to review and endorse during the workshop, these detailed grade- and content-area-specific descriptions communicate exactly what students performing at each level know and can do.

3. **Threshold PLDs.** Typically created during and used for standard setting only, these describe what a student just barely scoring into each performance level knows and can do. They may also be called Target PLDs or Just Barely PLDs.

4. **Reporting PLDs.** These are much-abbreviated PLDs (typically 350 or fewer characters) created following state approval of the performance standards used to describe student performance on score reports.

5.4.1.1 Science Range Performance-Level Descriptor Development

CAI and staff from participating states' Departments of Education (DOEs) reviewed existing range PLDs from several states' assessments based on three-dimensional science standards. States selected the range PLDs based on the standards drafted by the Washington State Office of Superintendent of Public Instruction (OSPI) as a starting point. Subsequently, CAI, state DOE staff, and educators from multiple states using science assessments based on the Shared Science Assessment Item Bank convened in May 2018 to review and refine the draft range PLDs.[4] The panels created policy PLDs and reviewed and identified refinements to the range PLDs to describe observable evidence for what student performance looks like in science at each performance level and test. CAI and one of the authors of the Next Generation Science Standards (NGSS) reviewed and applied recommendations to the PLDs. They ensured consistency, coherence, and articulation across tests and levels. Appendix 3-B, Development of Science Range Performance-Level Descriptors, provides additional information about the development of the range PLDs prior to states' standard-setting workshops.

5.4.1.2 OPI and Panelist Range Performance-Level Descriptor Review

OPI then reviewed the PLDs to ensure that the language accurately represented the goals and policies of the state. CAI worked with them to make revisions where necessary.

The day prior to the standard-setting workshop, the group of Montana educators selected to be standard-setting table leaders, who were intimately familiar with students and the subject matter, convened to review, revise, and approve the range PLDs. Appendix C, Montana Science Assessment Range Performance-Level Descriptors, provides the final range PLDs for the MSA.

5.4.2 Ordered Scoring Assertion Booklets

Like the Bookmark method used for establishing performance standards for traditional science tests, the AMP uses booklets of ordered test materials for setting standards. Instead of test items, the AMP uses scoring assertions presented in test-specific booklets called ordered scoring assertion booklets (OSABs). Each OSAB represents one possible testing instance resulting from applying the test blueprints to the state item pool.

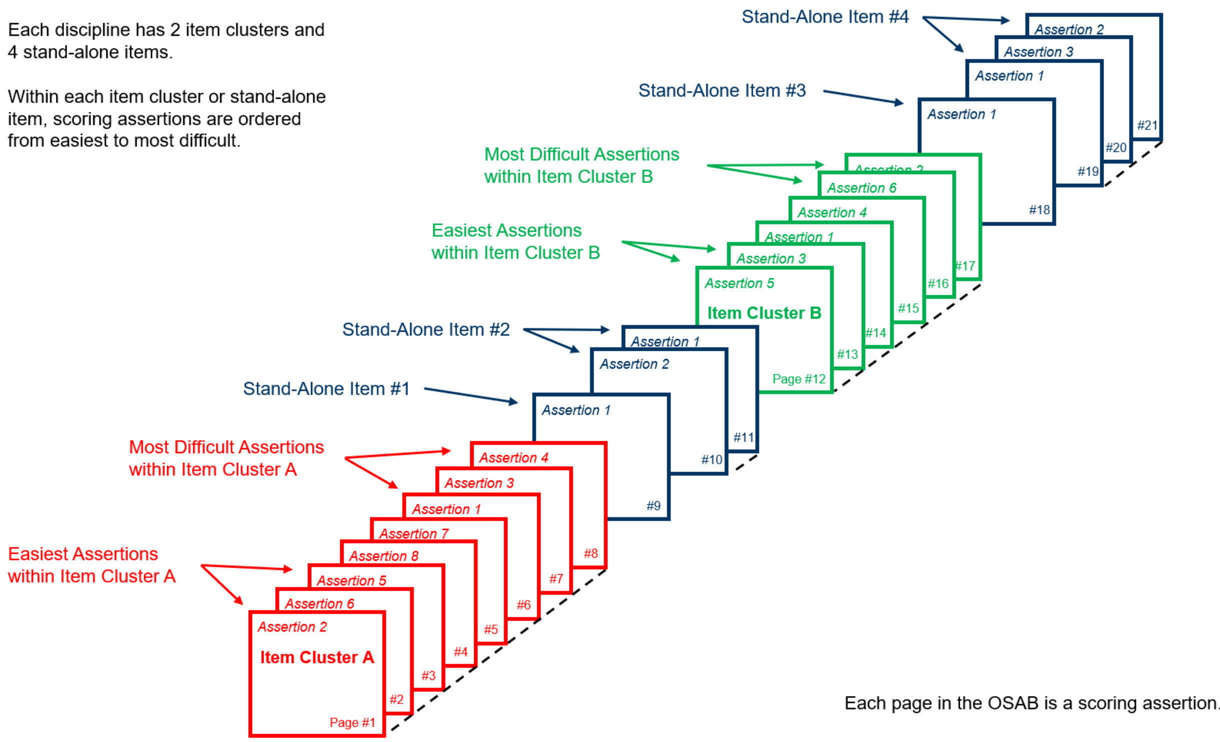
The OSABs were assembled using a mixed-integer programming approach. The objective function that was minimized was the number of gaps between the impact values of the assertions across the entire OSAB. A gap was defined as a difference of 3% or more between the impact values of two consecutive assertions ordered by difficulty. The linear constraints of the mixed-integer problem

[4] These states included Hawaii, New Hampshire, Oregon, Rhode Island, Utah, Vermont, West Virginia, and Wyoming.

represented the constraints implied by the blueprint. In addition, the total number of assertions was not allowed to exceed 85 to keep the OASB within a typical range, and to keep the burden on panelists reasonable. A set of feasible solutions was further evaluated based on the distribution of the impact values of assertions across the OSAB. The candidate solution was then reviewed internally by content experts and by the OPI. Two of the items in the grade 5 initial candidate solution were replaced by another two eligible items, upon OPI's request. The grade 8 candidate solution was approved without any changes.

Figure 7 below describes the structure of the OSAB.

Figure 7. Ordered Scoring Assertion Booklet (OSAB)



For the operational test, the order of the items was randomized over students. The items in the OSABs were grouped by science content area, so that panelists work through all items associated with one content area before moving on to the next, allowing panelists to focus on the knowledge and skill requirements for one content area at a time. For the grade 5 OSAB, the Life Sciences discipline items were presented first, then Physical Sciences items, and then Earth and Space Sciences items. For the grade 8 OSAB, the Physical Sciences discipline items were presented first, then Life Sciences items, and then Earth and Space Sciences items. Two item clusters and four stand-alone items represented each content area. Within a content area, the item clusters were presented first, followed by the stand-alone items. The item clusters and stand-alone items were further ordered by mean difficulty of the assertions within the item. This approach may help to reduce some of the cognitive demands on panelists by making clear that some items, and their associated interactions, are easier for students to access, even though the assertions they support are similar in content.

Within each item cluster or stand-alone item, scoring assertions were ordered by difficulty. Easier assertions are those that most students were able to demonstrate, and difficult assertions are those that the fewest students were able to demonstrate. Note that assertions were ordered by difficulty within items only. Across all items, this was generally not the case; for example, the most difficult assertion of an item presented early in the OSAB was typically more difficult than the easiest assertion of the next item in the OSAB. That is, the order of assertions in Figure 7. Ordered Scoring Assertion Booklet (OSAB) represents the order of presentation to the panelists, but assertions were not ordered by overall difficulty across all items (see Figure 8 for a depiction of the overlapping difficulty of assertions in the complete OSAB).

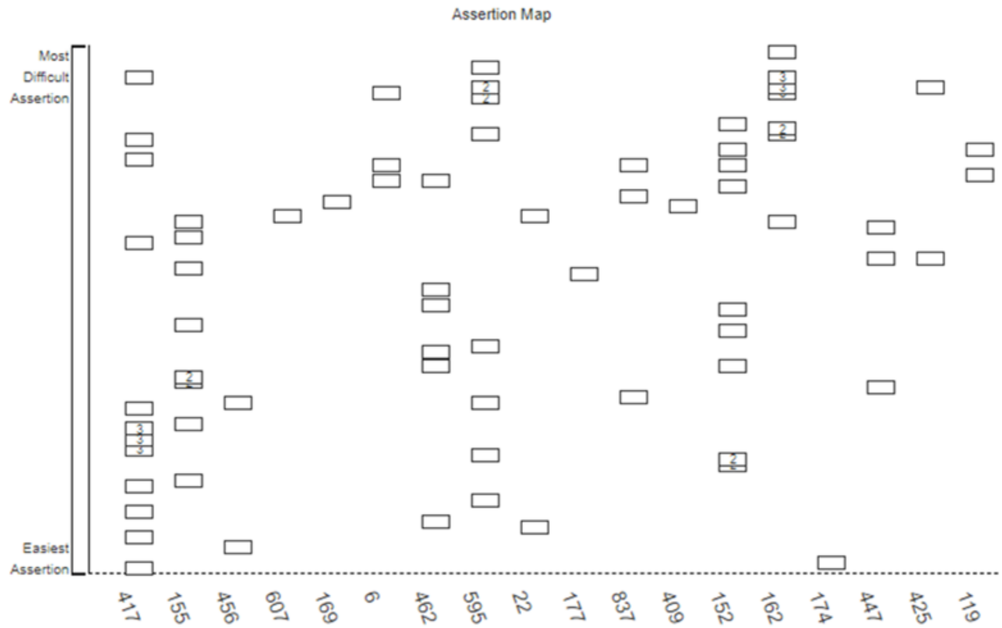
Not all items have assertions that will map onto all performance levels. For example, an item cluster may have assertions that map onto "Level 1," "Level 2," and "Level 3," but not "Level 4."

Each OSAB contains three content areas and 18 items (item clusters and stand-alone items). The grade 5 OSAB contained 72 assertions and the grade 8 OSAB contained 75 assertions. Each comprised six item clusters and 12 stand-alone items.

5.4.3 Assertion Maps

Assertion maps were provided to panelists to help reduce the cognitive load of the AMP. The assertion maps were displayed in CAI's online standard-setting tool and listed all scoring assertions in each OSAB by item ID, assertion, and plotted all assertions by difficulty. The assertion maps provided panelists with context about student performance on the assertions in the OSAB, describing the difficulty of each assertion in the underlying OSAB. This was to help panelists easily identify more- or less-difficult assertions and compare the difficulty of assertions across items. The assertion maps were provided during the OSAB review. After Round 1, the assertion maps were updated to also display the tentative standards (more details in Section 5.7.2.2, Feedback Data). Figure 8 presents the assertion map for grade 5. The assertion maps for both tests are presented in Appendix D, Standard-Setting Assertion Maps.

Figure 8. Standard-Setting Assertion Map, Science Grade 5



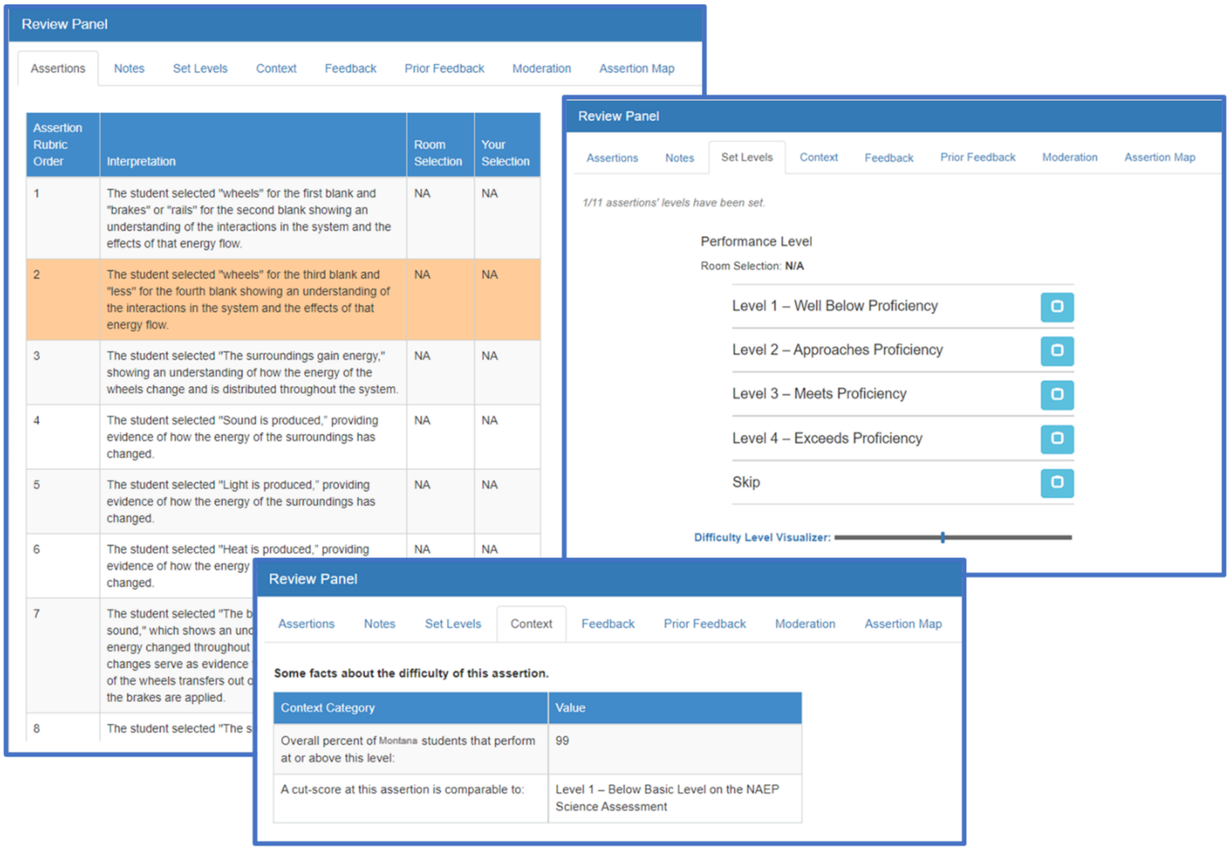
5.5 WORKSHOP TECHNOLOGY

The standard-setting panelists used CAI's online application for standard setting. All panelists used their own computers on which they took the test; reviewed item clusters, stand-alone items, and ancillary materials; and mapped assertions to performance levels.

Using tabs in the review panel of the tool (see Figure 9), panelists could review the items and scoring assertions; determine the relative difficulty of assertions to other assertions in the same item; examine the content alignment of each item (via the alignment of the assertions within an item, which all align to the same PE); assign assertions to performance levels; add notes and comments on the assertions as they reviewed them; and review contextual information and feedback data. Additionally, they had access to a difficulty level visualizer, a graphic representation of the difficulty of each assertion relative to all other assertions in the OSAB (not just within the item).[5] Panelists also reviewed their assertion placement, their table's placement, the other tables' placement, and the overall placement for all tables.

[5] The difficulty level visualizer represented the percentage of students whose ability level would fall at or above the difficulty level of that assertion.

Figure 9. Example Features in Standard-Setting Tool



Full-time CAI information technology specialists answered questions and ensured that technological processes ran smoothly and without interruption throughout the meeting.

5.6 Events

The standard-setting workshop occurred over a period of two days. Table 7 summarizes each day's events, and this section describes each event listed in greater detail. Appendix 3-E, Standard-Setting Workshop Agenda, provides the full workshop agenda.

Table 7. Standard-Setting Agenda Summary

Day 1: Tuesday, August 2, 2022
• Large-Group Orientation • Review and Take the Operational Test • Review Range PLDs • Discuss Threshold PLDs • OSAB Review
Day 2: Wednesday, August 3, 2022
• Continue OSAB Review • Assertion-Mapping Training • Round 1 Assertion Mapping

- Round 1 Feedback and Impact Data Review and Discussion
- Round 2 Assertion Mapping
- Round 2 Feedback and Impact Data Review
- Standard-Setting Workshop Evaluations
- Across-Grade Moderation and Articulation

5.6.1 Participant Login

Panelists were required to attend a technical check prior to the standard-setting workshop to ensure that they had access to the required sites needed to participate in the workshop. They also received and signed affidavits of nondisclosure at this time, affirming that they would not reveal any secure information they would have access to during the workshop. Panelists arrived at the workshop, virtually, on the first day, and followed the instructions given for joining the workshop via Microsoft Teams.

5.6.2 Large-Group Orientation

Ashley McGrath, Montana OPI Assessment Director, welcomed panelists to the workshop and provided context and background for the MSA. Ms. McGrath outlined the roles and responsibilities of the participants at the workshop: panelists, CAI staff, and OPI personnel. Dr. Stephan Ahadi then oriented participants to the workshop by describing the purpose and objectives of the meeting, explaining the process to be implemented to meet those objectives, and outlining the events that would happen each day. Dr. Ahadi explained that panelists were selected because they were experts, and how the process to be implemented over the two days was designed to elicit and apply their expertise to recommend new cut scores. Finally, he described how standard setting works and what would happen once the panelists had finalized their recommendations. Appendix F, Standard-Setting Training Slides, provides the slides used during the large-group training.

5.6.3 Confidentiality and Security

Workshop leaders and room facilitators addressed confidentiality and security during orientation and again in each room. Standard setting uses live science test items from the operational MSA, requiring confidentiality to maintain their security. Participants were not allowed to do the following either during, or after, the workshop:

- Discuss the test items outside of the meeting
- Discuss judgments or cut scores (their own or others') with anyone outside of the meeting
- Discuss secure materials with non-participants
- Create any form of electronic copy of test content (screen captures, electronic notes, etc.)
- Create any handwritten notes of test content
- Use a computer during the course of the meeting for any purpose other than participating in the standard-setting workshop and item review (e.g., email, web browsing, social media)
- Save notes about item or passage content to a computer

Participants could have general conversations regarding the process and days' events, but workshop leaders warned them against discussing details, particularly those involving test items, cut scores, and any other confidential information.

5.6.4 Take the Test

Following the large-group orientation, panelists broke out into their separate test-level virtual meeting rooms. As their introduction to the standard-setting process, panelists took a form of the test that students took in 2022, in the grade band to which they would be setting performance standards. They took the tests online via the same tool used to deliver operational tests to students, and the testing environment closely matched that of students when they took the test.

Taking the same test as students take provides the opportunity to interact with and become familiar with the test items and the look and feel of the student experience while testing. They could score their responses and had 90 minutes to interact with the test.

5.6.5 Range Performance-Level Descriptor Review

After taking the operational test, panelists completed a thorough review of the range PLDs for their assigned test. Panelists were provided with an overview of the PLDs and their importance to standard setting. The PLDs were used as a reference for evaluating student performance, so it was important for panelists to understand the critical role of PLDs in the standard-setting process.

Panelists began their review of the range PLDs that define what students in each performance level know and can do with respect to the Montana Science Content Standards. Workshop facilitators provided panelists with draft range PLDs, test blueprints, and the Montana Science Content Standards. The facilitators led panelists through a thorough review of the range PLDs for their assigned test using the materials as references and drawing on the expertise of the panelists.

Panelists identified key words describing the skills necessary for performance at each level and discussed the skills and knowledge that differentiate performance in each of the four levels.

Reviewing the range PLDs ensured that participants understood what students in Montana should know and be able to do and how much knowledge and skill students are expected to demonstrate at each level of performance.

5.6.6 Discuss Threshold Performance-Level Descriptors

After reviewing and discussing the range PLDs, panelists worked in their test-level groups to develop a shared understanding of the threshold PLDs that describe the skills that students just barely able to score in one performance level have but that students scoring just below the performance level do not have. Facilitators encouraged panelists to consider the characteristics of students who just barely qualify for entry into the performance level from those just below. Looking at each PLD, panelists identify the skills needed to just barely perform at that level. The following two questions guide the process:

1. What skills and knowledge must the student demonstrate to qualify for entrance into this performance level?
2. How does this differ from the upper range of the adjacent (lower) performance level?

These discussions yielded common descriptions of students just barely characterized by each PLD within each room.

The AMP employs the range PLDs since panelists are mapping items across the full range of the PLD. The purpose of the threshold PLD discussion was to enhance the panelists' understanding of the differences between PLD levels by paying special attention to the transition areas between performance levels.

5.6.7 Ordered Scoring Assertion Booklet Review

After reviewing and discussing the PLDs, panelists reviewed the item clusters, stand-alone items, and assertions in the OSAB. They took notes on each assertion to document the interactions required by each and described why an assertion might be more or less difficult than the previous assertion within the item. They also noted how each assertion related to the PLDs.

After reviewing the item interactions and scoring assertions individually, panelists engaged in discussion with group members about the skills required and relationships among the reviewed test materials and performance levels. This process ensured that panelists built a solid understanding of how the scoring assertions relate to the item interactions and how the items relate to the PLDs, and also helped to facilitate a common understanding among workshop panelists.

5.6.8 Assertion-Mapping Training

After reviewing the entire OSAB, facilitators described the processes for mapping assertions and determining cut scores. They explained that the objective of standard setting is aspirational; to identify what all students should know and be able to do, and not to describe what they currently know and can do.

Panelists were to match each assertion to the performance level best supported by the assertion using the PLDs; the difficulty level visualizer (described in Section 5.5, Workshop Technology); the assertion map (described in Section 5.4.3, Assertion Maps); their notes from the OSAB review; and their professional judgments. Figure 10 graphically describes the assertion-mapping process.

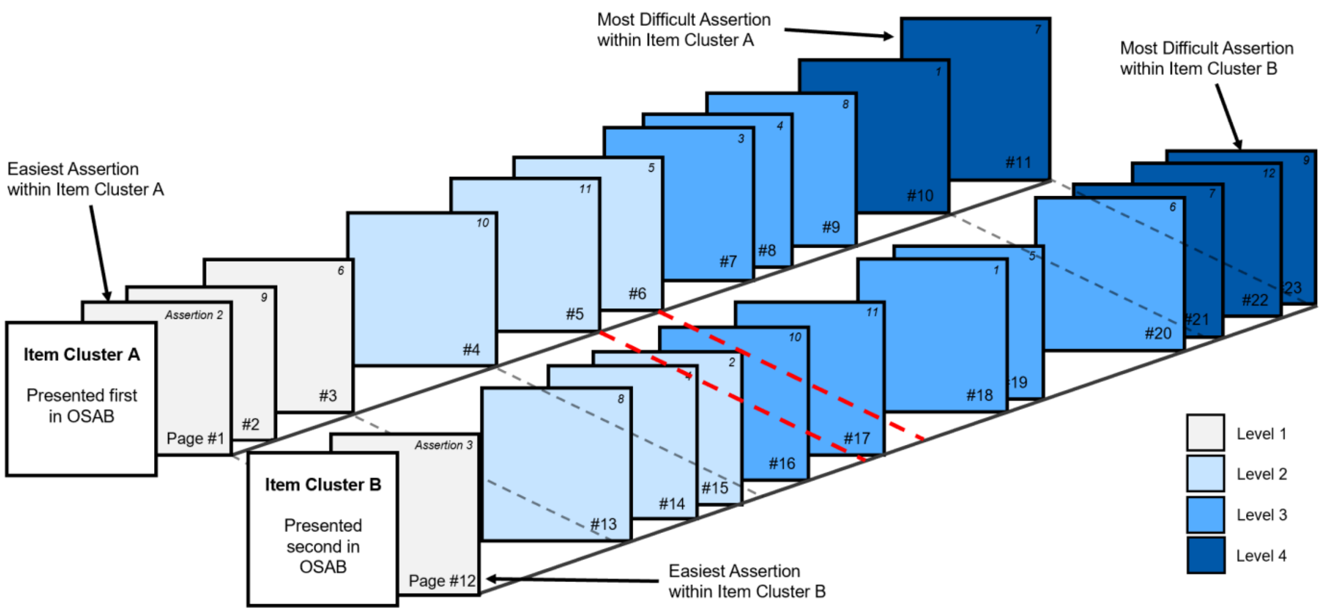
Facilitators provided the following process to guide the mapping of assertions onto PLDs:

1. How does the student interaction give rise to the assertion? Did they plot, select, or write something?
2. Why is this assertion more difficult to achieve than the previous one (within the item)?
3. Which PLD most ably describes this assertion and the underlying interactions?

It was emphasized that assertions within an item were ordered by difficulty, and therefore, the assigned performance levels should be ordered, as well. Within each item, panelists were not allowed to place an assertion into a lower performance level than the level at which the previous assertions had been placed. If panelists felt very strongly that an assertion was out of order in the OSAB, they were asked to skip (not assign any performance level to) the assertion. However, this was to be used as a last resort.

As assertion mapping was conducted independently for each item, there may not have been a flawless alignment of the assigned assertion levels across all items based on assertion difficulty. It was allowed (and it occurred frequently) that an assertion of one item had a higher difficulty but lower assigned performance level than another assertion from a different item (i.e., mapping inversions of assertions could occur across items, but mapping inversions of assertions were not allowed within an item). For example, in Figure 10, the difficulty of the assertion on page 6 of item cluster A ("Level 2") has a higher difficulty than the assertion on page 17 of item cluster B ("Level 3"). However, it was expected for the higher performance levels to be assigned more frequently with increasing assertion difficulty across items. Appendix 3-F, Standard-Setting Training Slides, provides the training slides used during the breakout room training.

Figure 10. Example of Assertion Mapping



Note. Figure 10 describes scoring assertion mapping across two item clusters, where the assertions on pages 1–3 and 12 are mapped onto Level 1; the assertions on pages 4–6 and 13–15 are mapped onto Level 2; the assertions on pages 7–9 and 16–20 are mapped onto Level 3; and the assertions on pages 10, 11, and 21–23 are mapped onto Level 4.

5.6.9 Practice Quiz

Panelists completed a practice quiz before beginning a practice round. The quiz assessed panelists' understanding in multiple ways. They must be able to perform the following:

- Describe where "Just Barely" students fall on a performance scale
- Indicate on a diagram how performance standards define performance levels
- Identify more- and less-difficult scoring assertions in the OSAB
- Answer questions about the assertion-mapping process and online application

Room facilitators reviewed the quizzes with the panelists and provided additional training for incorrect responses on the quiz. Appendix 3-G, Standard-Setting Practice Quiz, provides the quiz that panelists completed before mapping any assertions.

5.6.10 Practice Round

Following the practice quiz, panelists practiced mapping assertions to PLDs in a short practice OSAB consisting of one item cluster and one stand-alone item. The purpose of the practice round was to ensure that panelists were comfortable with the technology, items, item interactions, and scoring assertions before mapping any assertions in the OSAB. Panelists discussed their practice mappings and asked questions, and the room facilitators provided clarifications and further instructions until everyone had completed the practice round.

5.6.11 Readiness Form

After completing the practice round, and before mapping assertions to performance levels in Round 1, panelists completed a readiness assertion form. On this form, panelists asserted that their training was sufficient for them to understand the following concepts and tasks:

- The knowledge and skills described by the PLDs, and the skills and interactions that differentiate levels
- The structure, use, and importance of the OSAB
- The process to determine and map assertions to PLDs in the standard-setting tool
- Understanding how to use the assertion map when reviewing the OSAB and considering assertion-mapping decisions
- Understanding the contextual information (student impact data and benchmarking data) when considering assertion mapping decisions
- Readiness to begin the Round 1 task

The readiness form for Round 2 focused on affirming an understanding of the feedback data supplied after Round 1. On this form, all panelists affirmed the following:

- Understanding of the feedback data and impact data
- Understanding of the Round 2 task
- Readiness to complete the Round 2 task

Room facilitators reviewed the readiness forms and provided additional training to panelists not asserting understanding or readiness. However, every panelist affirmed readiness before mapping assertions in both rounds of the workshop. Appendix 3-H, Standard-Setting Readiness Forms, provides the forms that panelists completed prior to each round of standard setting.

5.7 ASSERTION MAPPING

Panelists mapped assertions independently, using the PLDs, their notes from reviewing each assertion, the difficulty level visualizer, the assertion map, and contextual information to place each of the assertions into one of the four performance levels.

5.7.1 Calculating Cut Scores from the Assertion Mapping

Cut scores were calculated by treating every possible scale value as a hypothetical cut score and evaluating the number of discrepancies between the assertion mappings of the panelists and the performance levels of the assertions implied by hypothetical cut score. The implied performance level of an assertion was determined by comparing the response probability of an assertion to the hypothetical cut.[6] Each cut score was defined as the score point that minimized the weighted number of discrepancies. The weights were defined as the inverse of the observed frequencies of each level. For each cut score, only the assertion mappings for the two adjacent levels were considered (e.g., for the second cut, only the assertions that were mapped onto "Approaches Proficiency" and "Level 3" were used). Specifically, let n_k be the number of assertions put at performance level k, t_k be the cut to be estimated, d_i be the assigned performance level, and θ_i be the response probability (RP) value of the *i*th assertion. For each assertion placed at levels k and $k+1$, the misclassification indicator is defined as

$$z_{ik}|t_k = \begin{cases} 1 \text{ if } (d_i = k \text{ and } t_k \leq \theta_i) \text{ or } (d_i = k+1 \text{ and } t_k > \theta_i) \\ 0 \text{ otherwise} \end{cases}$$

The cut t_k is then estimated by minimizing a loss function based on the weighted number of misclassifications

$$\underset{t_k}{\arg\min}\left(\frac{1}{n_k}\sum_{i\in\{d_i=k\}} z_{ik}|t_k + \frac{1}{n_{k+1}}\sum_{i\in\{d_i=k+1\}} z_{ik}|t_k\right)$$

Unlike the Bookmark method, the cut scores for a table or room were not the median value of the cut scores of the individual panelists. Instead, cut scores at the table and room (test) level were computed using the same method but considering the assigned levels of all the raters at the table and in the room, respectively. Applying these cut scores to the 2022 operational test data created data describing the percentage of students falling into each performance level. This algorithm calculated cut scores from the assertion mappings by panelist, table, and for the room.

[6] Typically, the response probability used in standard setting is .67 ("RP67" [Huynh, 1994]). RP67 is the assertion difficulty point where 67% of the students would earn the score point. The reason to adopt RP50 in some states, including the MSA, was because the difficulty of most items exceeded students' abilities. RP50 better aligned with the PLD and therefore led to more appropriate performance cut scores. Using RP50 prevented panelists from mapping the first cut score onto the lowest-difficulty assertions on the test. This approach has been adopted for other high-stakes tests, such as the Smarter Balanced Assessments (see Cizek & Koons, 2014).

5.7.2 Contextual Information and Feedback Data

To be adoptable, performance standards for a statewide system must be coherent across tests and subjects. They should be orderly across subjects with no dramatic differences in expectation. The following are characteristics of well-articulated standards:

- The cut scores for each performance level increase smoothly with each increasing grade.
- The cut scores should result in a reasonable percentage of students at each performance level; reasonableness can be determined by the percentage of students in the performance levels on historical tests, or contemporaneous tests measuring the same or similar content.
- Barring significant content standard changes (e.g., major changes in rigor), the percentage proficient on new tests should not be radically different from the percentage proficient on historical tests.

The standard-setting tool developed by CAI provides feedback data and allows for displaying contextual information to ensure that standard-setting recommendations are well articulated.

5.7.2.1 Contextual Information

During OSAB review, panelists were also provided with additional contextual information to help inform their primary content-driven performance standard recommendations. The standard-setting tool developed by CAI allows for displaying both impact and benchmark data to ensure that standard-setting recommendations are well articulated. The contextual information provided included impact data and benchmark data for each of the assertions of the OSAB, as described in the following sections.

Impact Data

The impact data for an assertion was defined as the percentage of students who performed at or above the specified RP value associated with the assertion. Panelists were asked to consider the impact data when making their content-based assertion mappings.

Benchmark Data

The National Assessment of Educational Progress (NAEP) 2015 Montana science performance level percentages (NAEP, 2015) provided benchmark data, another source of contextual information that panelists could use to evaluate and adjust their assertion mapping. By comparing the results of each round against the percentage proficient in Montana on NAEP science, panelists could evaluate the reasonableness of the proposed performance standards. NAEP provides national-level data in science for grades 4 and 8; benchmark data for grade 5 were interpolated. For each ordered scoring assertion, panelists were provided with the associated performance level for the NAEP science. An example of the benchmark information provided for each assertion in the review panel of the standard-setting tool is shown in Figure 9. The 2015 NAEP benchmark data were also graphically shown on the left side of the assertion map (see Figure 12).

5.7.2.2 Feedback Data

The online standard-setting tool created feedback data and cut scores corresponding to the assertion mappings for each panelist, for each table, and for the room overall (across both tables). In addition, panelists were shown impact data based on the cut scores resulting from their assertion mappings. Impact data were defined for panelists as the percentages of students who would reach or exceed each of the performance standards given the assertion mappings. Percentages were calculated using the student data from the 2022 MSA test administration. This information allowed panelists to compare their mappings to other panelist's mappings to evaluate the impact of their current mappings.

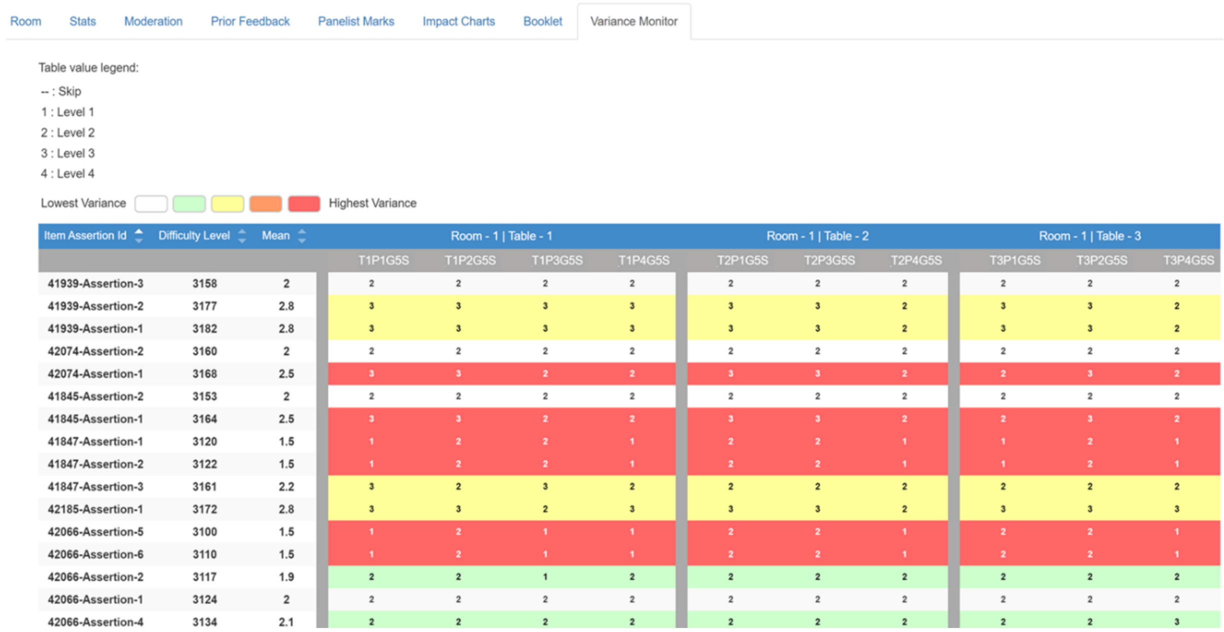
The standard-setting tool also generated variance monitor data and the assertion maps in the tool were updated to display the tentative standards for panelists to evaluate before Round 2 (the variance data and assertion maps are described in more detail in the following paragraphs). All feedback and information served to inform, but not determine, their Round 2 decisions. Panelists discussed this information and the impact that the Round 1 cut scores may have on students before mapping assertions in Round 2.

After reviewing the feedback data, the workshop facilitators provided panelists with additional instructions for completing Round 2. First, they described the goal of Round 2 as one of convergence, but not consensus, on a common performance standard. The second goal was to encourage articulation across grade levels. Each room spent time reviewing and discussing assertion mappings and articulation. After completing these discussions, panelists again worked through mapping all OSAB assertions to performance levels for Round 2.

Variance Monitor Data

Feedback included a review of a variance monitor, part of CAI's online standard-setting tool that color codes the variance of assertion classifications. For all assertions, the variance monitor shows the performance level to which each panelist assigned the assertion. The tool highlights assertions that panelists have assigned to different performance levels. Figure 11 illustrates the types of information available in the variance monitor. Room facilitators and panelists reviewed and discussed the assertions with the most variable mappings.

Figure 11. Variance Monitor in CAI's Standard-Setting Tool

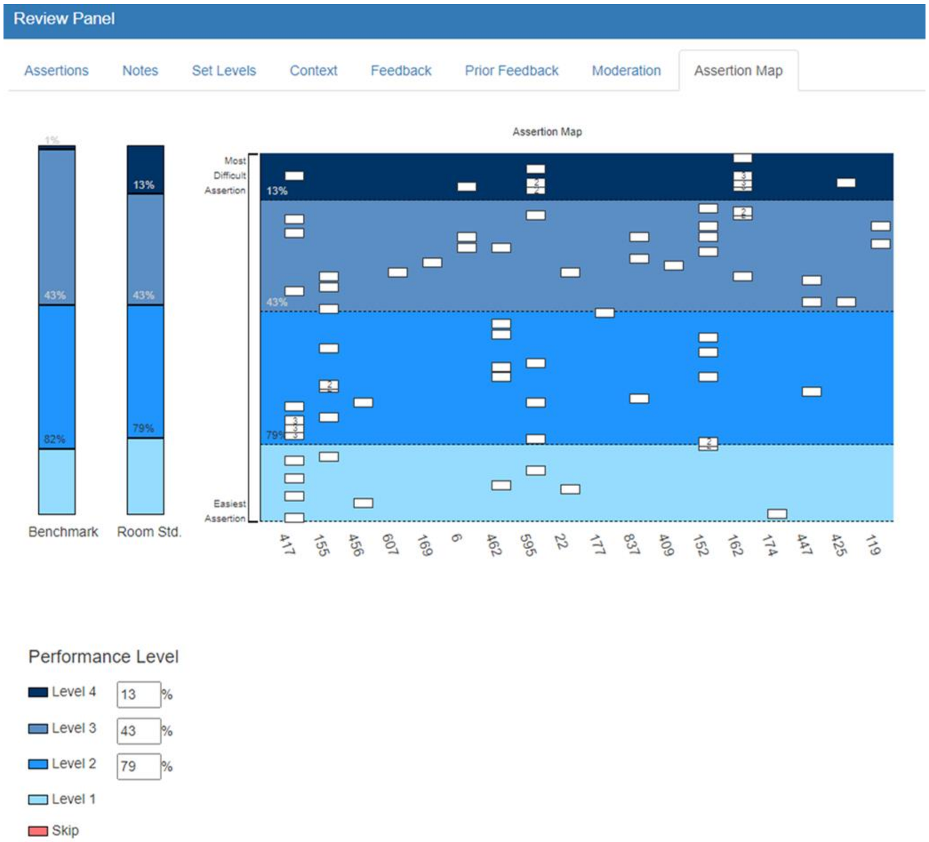


Item Assertion Id	Difficulty Level	Mean	Room - 1 | Table - 1				Room - 1 | Table - 2			Room - 1 | Table - 3		
			T1P1G5S	T1P2G5S	T1P3G5S	T1P4G5S	T2P1G5S	T2P3G5S	T2P4G5S	T3P1G5S	T3P2G5S	T3P4G5S
41939-Assertion-3	3158	2	2	2	2	2	2	2	2	2	2	2
41939-Assertion-2	3177	2.8	3	3	3	3	3	3	2	3	3	2
41939-Assertion-1	3182	2.8	3	3	3	3	3	3	2	3	3	2
42074-Assertion-2	3160	2	2	2	2	2	2	2	2	2	2	2
42074-Assertion-1	3168	2.5	3	3	2	2	3	3	2	2	3	2
41845-Assertion-2	3153	2	2	2	2	2	2	2	2	2	2	2
41845-Assertion-1	3164	2.5	3	3	2	2	3	3	2	2	3	2
41847-Assertion-1	3120	1.5	1	2	2	1	2	2	1	1	2	1
41847-Assertion-2	3122	1.5	1	2	2	1	2	2	1	1	2	1
41847-Assertion-3	3161	2.2	3	2	3	2	2	2	2	2	2	2
42185-Assertion-1	3172	2.8	3	3	2	3	3	3	2	3	3	3
42066-Assertion-5	3100	1.5	1	2	1	1	2	2	1	2	2	1
42066-Assertion-6	3110	1.5	1	2	1	1	2	2	1	2	2	1
42066-Assertion-2	3117	1.9	2	2	1	2	2	2	2	2	2	2
42066-Assertion-1	3124	2	2	2	2	2	2	2	2	2	2	2
42066-Assertion-4	3134	2.1	2	2	2	2	2	2	2	2	2	3

Assertion Maps

In addition to providing the numerical value of the cut scores and impact data, the feedback was also shown on the assertion maps. After Round 1, the assertion maps displayed in CAI's online standard-setting tool are updated with the overall room cut scores and the individual panelist cut scores for Round 1. Figure 12 presents the assertion map for grade 5 with the overall room cut scores for Round 1. The Round 1 and Round 2 assertion maps with overall room cut scores for all three tests are presented in Appendix I, Round 1 and Round 2 Standard-Setting Assertion Maps.

Figure 12. Round 1 Standard-Setting Assertion Map, Grade 5



Panelists were instructed to consider their assertion mappings to compare the room cut score and assertions to their cut scores and assertion mappings. They were again reminded to evaluate the relative location of the assertions on the assertion maps.

5.8 Assertion Mapping Results

The CAI online standard-setting tool automatically computes the results and impact data for each
5.8.1 round and then CAI room facilitators and psychometricians present the Round 1 results and feedback data for each test.

Round 1 Results

Table 8 presents the performance standards and associated impact data (percentage of students falling at or above each of the performance standards based on the recommended Round 1 cut scores) from Round 1.

Table 8. Round 1 Results

Test and Table	Cut Score			Impact Data		
	Level 2	*Level 3*	*Level 4*	*Level 2*	*Level 3*	*Level 4*
Grade 5	478	506	531	79	43	13

Test and Table	Cut Score			Impact Data		
	Level 2	*Level 3*	*Level 4*	*Level 2*	*Level 3*	*Level 4*
Table 1	478	504	NA	79	46	NA
Table 2	477	506	531	80	43	13
Grade 8	781	808	826	76	40	15
Table 1	781	808	826	76	40	15
Table 2	781	808	840	76	40	7

Note. The test row summarizes the room data (across both tables). Impact data describes the percentage of students falling at or above each of the performance standards based on the recommended Round 1 cut scores.

Reviewing the Round 1 results began with a discussion of the feedback data from Round 1, beginning with table-level feedback and discussion, progressing to the room-level discussion. After reviewing the feedback (i.e., individual cuts, cuts by a table, cuts by a room) and impact data, workshop facilitators provided panelists with additional instructions for completing Round 2. They described the goal of Round 2 as one of convergence, but not consensus on a common performance standard. The room then spent time reviewing and discussing assertion mappings. After completing these discussions, panelists again worked through the OSAB, mapping assertions for Round 2.

5.8.2 Round 2 Results

Table 9 presents the recommended performance standards and associated impact data (percentage of students falling at or above each of the performance standards based on the recommended Round 2 cut scores) for Round 2.

Table 9. Round 2 Results

Test and Table	Cut Score			Impact Data		
	Level 2	*Level 3*	*Level 4*	*Level 2*	*Level 3*	*Level 4*
Grade 5	477	506	531	80	43	13
Table 1	477	504	531	80	46	13
Table 2	477	506	531	80	43	13
Grade 8	781	808	826	76	40	15
Table 1	778	808	826	79	40	15
Table 2	781	808	826	76	40	15

Note. The test row summarizes the room data (across both tables). Impact data describes the percentage of students falling at or above each of the performance standards based on the recommended Round 2 cut scores.

Figure 13 represents those values graphically.

Figure 13. Percentage of Students Reaching or Exceeding Each Recommended Science Performance Standard in 2022

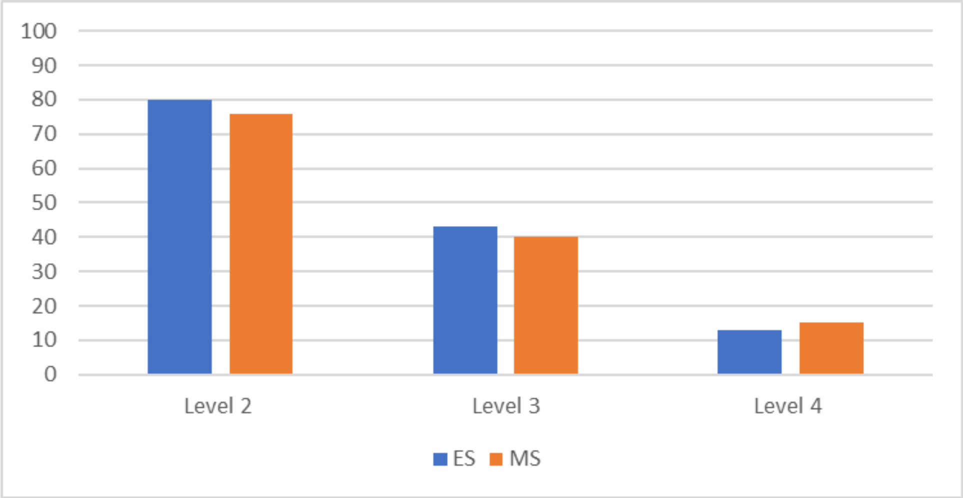
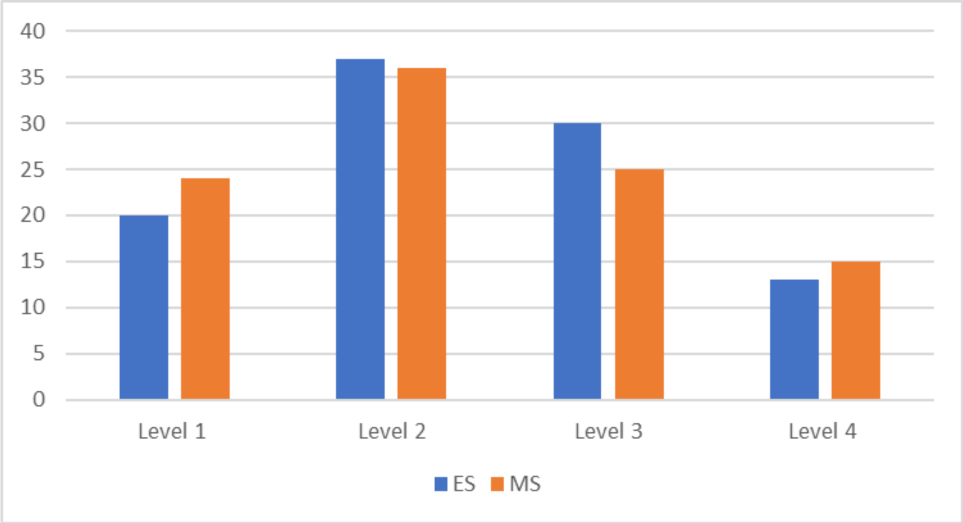


Table 10 indicates the percentage of students classified within each of the performance levels in 2022. The values are displayed graphically in Figure 14.

Table 10. Percentage of Students Classified Within Each Recommended Science Performance Level in 2022

Test	Level 1	Level 2	Level 3	Level 4
Grade 5	27	36	27	10
Grade 8	35	32	26	7

Figure 14. Percentage of Students Classified Within Each Recommended Science Performance Level in 2022



5.8.3 Convergence across Rounds

While consensus is not an objective of standard setting, convergence is. Indicators of panelist convergence over rounds are the interquartile range (IQR) and standard deviation (SD) of the standards computed for individual panelists based on their mappings. The IQR and SD for each test and after each round are presented in Table 11. For the Level 3 and Level 4 standards, the indicators consistently show that there is an improved convergence in individual standards over the rounds. For the Level 2 standard, although the IQR and SD slightly increased over the rounds, the values remained relatively small, indicating a good convergence in both rounds.

Table 11. Interquartile Range and Standard Deviation of Panelist-Recommended Performance Standards

Test	Statistic	Level 2 Approaches Proficiency		Level 3 Meets Proficiency		Level 4 Exceeds Proficiency	
		Round 1	*Round 2*	*Round 1*	*Round 2*	*Round 1*	*Round 2*
Grade 5	IQR	3.00	6.00	3.75	2.25	12.50	7.50
	SD	7.40	8.30	6.91	5.46	13.19	7.79
Grade 8	IQR	5.00	8.00	7.50	1.00	20.50	14.00
	SD	3.69	4.03	4.34	2.53	13.68	11.28

5.8.4 Moderation

Panelists receive the information necessary for articulation prior to Round 2. Often, panelists intuitively create well-articulated sets of performance standards, but sometimes minor changes might significantly improve articulation. OPI saw no need for moderation of the Round 2 recommended performance standards during the moderation session.

5.9 Workshop Evaluations

After finishing all activities, panelists completed online workshop evaluations independently, in which they described and evaluated their experience taking part in the standard setting. Table 12–17 summarize the results of the evaluations. Evaluation items endorsed by fewer than 90% of panelists are discussed in the text, and the least-endorsed items are discussed in terms of the number and type of response.

Workshop participants overwhelmingly indicated clarity in the instructions, materials, data, and process (see Table 12).

Table 12. Evaluation Results: Clarity of Materials and Process

Please rate the clarity of the following components of the workshop.	**Percentage Indicating "Somewhat Clear" or "Very Clear"**		
	Science Grade 5	*Science Grade 8*	*Overall*
Instructions provided by the workshop leader	100%	100%	100%
PLDs	100%	100%	100%
OSAB	100%	100%	100%
Assertion Map	100%	100%	100%
Impact Data (percentage of students that would achieve at the level indicated by the assertion difficulty)	100%	94%	97%
Panelist Agreement Data	100%	94%	97%

Note. Number of responses = 30 (grade 5 responses = 14, grade 8 responses = 16). Evaluation response options included "Very Unclear," "Somewhat Unclear," "Somewhat Clear," and "Very Clear."

As shown in Table 13, most panelists felt that the time allocated to various workshop tasks was about right, though a few panelists had suggestions regarding time allocation:

- Nine (three from grade 5 and six from grade 8) panelists reported that the large-group orientation was too long.
- Seven (one from grade 5 and six from grade 8) panelists indicated not having enough time for taking the test.
- Two panelists from grade 8 indicated having too little time to review the PLDs, while four grade 8 panelists reporting having too much time for this review.
- Two panelists from grade 8 reported having too little to discuss the skills demonstrated by students who are "Just Barely" described by each PLD, while two grade 8 panelists reported having too much time.
- Two panelists from grade 8 indicated that not enough time was allowed for reviewing the OSAB, while five other grade 8 panelists reported the review took too much time.

Table 13. Evaluation Results: Appropriateness of Process

How appropriate was the amount of time you were given to complete the following components of the standard-setting process?	**Percentage Indicating "About Right"**		
	Science Grade 5	*Science Grade 8*	*Overall*
Large-group orientation	79%	63%	70%
Experiencing the online assessment	93%	63%	77%
Reviewing the PLDs	100%	63%	80%
Discussion of the skills demonstrated by students who are "Just Barely" described by each PLD	100%	75%	87%
Reviewing the OSAB	100%	56%	77%

How appropriate was the amount of time you were given to complete the following components of the standard-setting process?	Percentage Indicating "About Right"		
	Science Grade 5	*Science Grade 8*	*Overall*
Mapping your scoring assertions to performance levels in each round	100%	100%	100%
Round 1 results discussion	100%	94%	97%

Note. Number of responses = 30 (grade 5 responses = 14, grade 8 responses = 16). Evaluation response options included "Too Little," "Too Much," and "About Right."

Participants appreciated the importance of the multiple factors contributing to assertion mapping, with nearly all participants rating each factor as important or very important (see Table 14).

Table 14. Evaluation Results: Importance of Materials

How important were each of the following factors in your mapping of scoring assertions to performance levels?	Percentage Indicating "Somewhat Important" or "Very Important"		
	Science Grade 5	*Science Grade 8*	*Overall*
PLDs	100%	100%	100%
"Just Barely" PLDs	100%	100%	100%
Your perception of the difficulty of the scoring assertions and items in general	100%	100%	100%
Your experience with students	100%	94%	97%
Discussions with other panelists	100%	100%	100%
Assertion map	93%	100%	97%
External benchmark data	100%	94%	97%
Impact data (percentage of students that would achieve at the level indicated by the assertion difficulty)	100%	100%	100%
Room-agreement data (room, table, and individual standards)	100%	100%	100%

Note. Number of responses = 30 (grade 5 responses = 14, grade 8 responses = 16). Evaluation response options included "Not Important," "Somewhat Important," and "Very Important."

Participant understanding of the workshop processes and tasks was consistently high (see Table 15). The one exception was that two grade 8 panelists disagreed with the statement regarding comfortability expressing their opinions throughout the workshop.

Table 15. Evaluation Results: Understanding Processes and Tasks

At the end of the workshop, please rate your agreement with the following statements.	Percentage Indicating "Agree" or "Strongly Agree"		
	Science Grade 5	*Science Grade 8*	*Overall*
I understood the purpose of this standard-setting workshop.	100%	100%	100%
	100%	100%	100%

At the end of the workshop, please rate your agreement with the following statements.	**Percentage Indicating "Agree" or "Strongly Agree"**		
	Science Grade 5	*Science Grade 8*	*Overall*
The procedures used to recommend performance standards were fair and unbiased.			
The training provided me with the information I needed to recommend performance standards.	100%	100%	100%
Taking the online assessment helped me to better understand what students need to know and be able to do to answer each assertion correctly.	100%	100%	100%
The PLDs (descriptions of what students within each performance level are expected to know and be able to do) provided a clear picture of expectations for student performance at each level.	100%	100%	100%
I was able to develop an understanding of the knowledge and skills demonstrated by students who are "Just Barely" described by the PLDs.	93%	100%	97%
I understood how to review each assertion in the OSAB to determine what students must know and be able to do to answer each assertion correctly.	100%	100%	100%
I understood how to map assertions to the most apt performance level.	100%	100%	100%
I found the assertion map helpful in my decisions about the assertions I mapped to performance levels.	93%	100%	97%
I found the benchmark data and discussions helpful in my decisions about the assertions I mapped to performance levels.	100%	94%	97%
I found the impact data (percentage of students that would achieve at the level indicated by the assertion difficulty) helpful when mapping assertions to performance levels.	100%	94%	97%
I found the panelist-agreement data (room, table, and individual standards) and discussions helpful when mapping assertions to performance levels.	93%	100%	97%
I felt comfortable expressing my opinions throughout the workshop.	100%	88%	93%
Everyone was given the opportunity to express his or her opinions throughout the workshop.	100%	100%	100%

Note. Number of responses = 30 (grade 5 responses = 14, grade 8 responses = 16). Evaluation response options included "Strongly Disagree," "Disagree," "Agree," and "Strongly Agree."

The majority of participants agreed that the standards set during the workshop reflected the intended grade-level expectations (see Table 16). However, five grade 5 panelists and four grade 8 panelists disagreed that students performing at Level 2 perform below expectations for the grade.

Table 16. Evaluation Results: Student Expectations

Please read the following statement carefully and indicate your response.	Percentage Indicating "Agree" or "Strongly Agree"		
	Science Grade 5	*Science Grade 8*	*Overall*
A student performing at "Level 2" performs below expectations for the grade.	64%	75%	70%
A student performing at "Level 3" meets expectations for the grade.	100%	100%	100%
A student performing at "Level 4" exceeds expectations for the grade.	100%	100%	100%

Note. Number of responses = 30 (grade 5 responses = 14, grade 8 responses = 16). Evaluation response options included "Strongly Disagree," "Disagree," "Agree," and "Strongly Agree."

5.9.1 Workshop Participant Feedback

Finally, panelists responded to two open-ended questions: "What suggestions do you have to improve the training or standard-setting process?" and "Do you have any additional comments? Please be specific."

Twenty-six panelists responded to the first question, and 21 responded to the second. Most responses indicated the training was effective and the process was clear. Participants provided minor suggestions, such as shortening or lengthening the time allocated for some tasks, holding the workshop in-person with opportunities for round table discussion before convening large group discussions, and providing more clarity and opportunity to work with terminology that panelists may not have any prior experience with. Many appreciated the organization, well-prepared materials, and technology, and many panelists complimented the professionalism and expertise of the facilitators.

Additional participant comments included:

"I really enjoyed this workshop and feel like I have a better understanding of how to not only prepare my students, but how to evaluate and assess their skills and content knowledge."

"This would be a great workshop for new teachers or first year teachers. It would be a great way to develop a deeper understanding of the standards, address those 'just barely' students and network with other colleagues."

"I've enjoyed working on this project and hope to continue to provide my time, even though I am retired. I still want to keep learning and helping students and teachers."

6. Validity Evidence

Validity evidence for standard setting is established in multiple ways. First, standard setting should adhere to the standards established by appropriate professional organizations and be consistent with the recommendations for best practices in the literature and established validity criteria. Second, the process should provide the evidence required of states to meet federal peer review requirements. We describe each of these in the following sections.

6.1 Evidence of Adherence to Professional Standards and Best Practices

The Montana Sciences Assessment (MSA) standard-setting workshop was designed and executed consistent with established practices and best-practice principles (Hambleton & Pitoniak, 2006; Hambleton, Pitoniak, & Copella, 2012; Kane, 2001; Mehrens, 1995). The process also adhered to the following professional standards recommended in the *Standards for Educational and Psychological Testing* (AERA, APA, & NCME, 2014) related to standard setting:

> *Standard 5.21:* When proposed score interpretations involve one or more cut scores, the rationale and procedures used for establishing cut scores should be documented clearly.
>
> *Standard 5.22:* When cut scores defining pass-fail or proficiency levels are based on direct judgments about the adequacy of item or test performances, the judgmental process should be designed so that the participants providing the judgments can bring their knowledge and experience to bear in a reasonable way.
>
> *Standard 5.23:* When feasible and appropriate, cut scores defining categories with distinct substantive interpretations should be informed by sound empirical data concerning the relation of test performance to the relevant criteria.

The sections of this documentation discussing the rationale and procedures used in the standard-setting workshop address Standard 5.21. The Assertion-Mapping Procedure (AMP) standard-setting procedure is appropriate for tests of this type—with interrelated sets of three-dimensional item clusters and scaled using item response theory (IRT). Section 5.1, The Assertion-Mapping Procedure, provides the justification for and the additional benefits of selecting the AMP to establish the cut scores. Section 5.6, Events, through Section 5.9.1, Round 1 Results, document the process followed to implement the method.

The design and implementation of the AMP address Standard 5.22. The method directly leverages the subject-matter expertise of the panelists placing assertions into performance levels and incorporates multiple, iterative rounds of ratings in which panelists modify their judgments based on feedback and discussion. Panelists apply their expertise in multiple ways throughout the process by

- understanding the test, test items, and scoring assertions (from an educator and student perspective);
- describing the knowledge and skills measured by the test;
- identifying the skills associated with each test scoring assertion;
- describing the skills associated with student performance at each performance level;
- identifying which test scoring assertions students at each performance level should be able to answer correctly; and
- evaluating and applying feedback and reference data to the Round 2 recommendations and considering the impact of the recommended cut scores on students.

Panelists' understanding of the AMP was assessed with a quiz before the practice round. Additionally, panelists' readiness evaluations provided evidence of a successful orientation to the process and understanding of the process, while their workshop evaluations provided evidence of confidence in the process and resulting recommendations.

The recruitment process resulted in panels that were representative of important regional and demographic groups who were knowledgeable about the subject area and students' developmental level. Section 5.3.4, Educator Participants, summarizes details about the panel demographics and qualifications.

The provision of benchmark, context, and articulation data to panelists after Round 1 addresses Standard 5.23 (see Section 5.7.2, Contextual Information and Feedback Data). This set of empirical data provides necessary and additional context describing student performance given the recommended standards.

6.2 Evidence in Terms of Peer Review Critical Elements

The United States Department of Education (ED) guides the peer review of state assessment systems. This guidance is intended to support states in meeting statutory and regulatory requirements under Title I of the Elementary and Secondary Education Act of 1965 (ESEA; U.S. Department of Education, 2015). The following critical elements are relevant to standard setting; evidence supporting each element immediately follows.

> **Critical Element1.5:** Meaningful consultation in the development of challenging state standards and assessments.

Montana educators played a critical role in establishing performance levels for the tests. They created the item clusters; reviewed and revised the Performance-Level Descriptors (PLDs); mapped assertions to performance levels to delineate performance at each performance level; considered benchmark data and the impact of their recommendations; and formally recommended performance standards.

Many subject-matter experts contributed to developing Montana's performance standards. Contributing educators were subject-matter experts in their content area, in the content standards and curriculum that they teach, and in the developmental and cognitive capabilities of their students. Cambium Assessment, Inc.'s (CAI) facilitators were subject-matter experts in the subjects tested and in facilitating effective standard-setting workshops. The psychometricians performing the analyses and calculations throughout the meeting were subject-matter experts in the measurement and statistics principles required of the standard-setting process.

> **Critical Element 6.2:** Achievement standards setting. The state used a technically sound method and process that involved panelists with appropriate experience and expertise for setting its academic performance standards.

Evidence to support this critical element includes:

1. The rationale for and technical sufficiency of the AMP selected to establish performance standards (Section 5.1, The Assertion-Mapping Procedure).

2. Documentation that the method used for setting cut scores allowed panelists to apply their knowledge and experience reasonably and supported the establishment of reasonable and defensible cut scores (Section 5.6, Events; Section 5.6.2, Large-Group Orientation; Section 5.9, Assertion Mapping Results; and Section 6.1, Evidence of Adherence to Professional Standards and Best Practices).

3. Panelists self-reported readiness to undertake the task (Section 5.6.9, Practice Quiz; and Section 5.6.11, Readiness) and confidence in the workshop process and outcomes (Section 5.9, Assertion Mapping Results; and Section 5.9.1, Round 1 Results) supporting the validity of the process.

4. The standard-setting panels consisted of panelists with appropriate experience and expertise, including content experts with experience teaching Montana's science content standards, and individuals with experience and expertise teaching special population and general education students in Montana (Section 5.3.4, Educator Participants; and Appendix 3-A, Standard-Setting Panelist Characteristics).

7. References

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

Bradlow, E. T., Wainer, H., & Wang, X. (1999). A Bayesian random effects model for testlets. *Psychometrika*, *64*(2), 153–168.

Cizek, G. J., & Koons, H. (2014). Observation and Report on Smarter Balanced Standard Setting: October 12–20, 2014. Retrieved from https://portal.smarterbalanced.org/library/en/standard-setting-observation-and-report.pdf.

Egan, K. L., Schneider, M. C., & Ferrara, S. (2012). Performance level descriptors: History, practice, and a proposed framework. In G. J. Cizek (Ed.) *Setting performance standards: Foundations, methods, innovations* (2nd ed., pp. 79–106). Routledge.

Ferrara, S., & Lewis, D. M. (2012). The item-descriptor (ID) matching method. In G. J. Cizek (Ed.), *Setting performance standards. Foundations, methods, and innovations* (2nd ed., pp. 255–282). New York: Routledge.

Gibbons, R. D., & Hedeker, D. R. (1992). Full-information bi-factor analysis. *Psychometrika*, *57*(3), 423–436.

Hambleton, R. K., & Pitoniak, M. J. (2006). Setting performance standards. In R. L. Brennan (Ed.), *Educational measurement* (4th ed., pp. 433–470). Westport, CT: Praeger.

Hambleton, R. K., Pitoniak, M. J., & Copella, J. M. (2012). Essential steps in setting performance standards on educational tests and strategies for assessing the reliability of results. In G. J. Cizek (Ed.), *Setting performance standards: Foundations, methods, and innovations* (2nd ed., pp. 47–76). New York: Routledge.

Huynh, H. (1994, Oct.). Some technical aspects in standard setting. In *Proceedings of the Joint Conference on Standard Setting for Large Scale Assessment Programs* (co-sponsored by National Assessment Governing Board and National Center for Education Statistics), Washington, DC, October 5–7, 1994, pp. 75–91.

Kane, M. T. (2001). So much remains the same: Conception and status of validation in setting standards. In G. J. Cizek (Ed.), *Setting performance standards: Concepts, methods, and perspectives* (pp. 53–88). Mahwah, NJ: Lawrence Erlbaum.

Kingston, N. M., Kahl, S. R., Sweeney, K. P., & Bay, L. (2001). Setting performance standards using the body of work method. In G. J. Cizek (Ed.), *Setting performance standards: Concepts, methods, and perspectives* (pp. 219–248). Mahwah, NJ: Lawrence Erlbaum Associates.

Mehrens, W. (1995). Legal and professional bases for licensure testing. In J. C. Impara (Ed.), *Licensure testing: Purposes, procedures, and practices*. Lincoln, NE: Buros Institute of Mental Measurements, University of Nebraska–Lincoln.

Mitzel, H. C., Lewis, D. M., Patz, R. J., & Greene, D. R. (2001). The Bookmark procedure: Psychological perspectives. In G. Cizek (Ed.), *Setting performance standards: Concepts, methods, and perspectives* (pp. 249-281). Mahwah, NJ: Lawrence Erlbaum Associates.

National Center for Education Statistics. (2015). *The Nation's Report Card*. National Assessment of Educational Progress. Retrieved from: https://nces.ed.gov/nationsreportcard/.

National Research Council. (2012). *A framework for K–12 science education: Practices, crosscutting concepts, and core ideas*. Washington, DC: The National Academies Press.

Rijmen, F. (2010). Formal relations and an empirical comparison among the bi-factor, the testlet, and a second-order multidimensional IRT model. *Journal of Educational Measurement*, *47*, 361−372.

Rijmen, F., Cohen, J., Butcher, T., & Farley, D. (2018, June 28). Scoring and reporting for assessments developed for the new science standards [Symposium]. National Conference on Student Assessment, San Diego, CA.Skaggs, G., Hein, S. F., & Awuor, R. (2007). Setting passing scores on passage-based tests: A comparison of traditional and single-passage bookmark methods. *Applied Measurement in Education*, *20*, 405–426.

U. S. Department of Education. (2015). *Non-Regulatory Guidance for States for Meeting Requirements of the Elementary and Secondary Education Act of 1965*. Washington, D.C. Retrieved from https://www2.ed.gov/policy/elsec/guid/assessguid15.pdf.

Appendix 3-A

Standard-Setting Panelist Characteristics

Standard-Setting Panelist Characteristics

Table A-1. Standard-Setting Panelists, Science Grade 5

Position	Location of Current Position	Gender	Race/ Ethnicity	Highest Level of Education	Years Teaching Experience	Years Professional Experience	Years Teaching/ Implementing the Montana Science Content Standards	School District Size	School District Urbanicity	Table Leader
General Education Teacher	School	Female	White	Master's degree (e.g., M.A., M.S.)	11 to 15 years	None	10	Large	Urban	Yes
Retired General Education Teacher	Other	Female	White	Master's degree (e.g., M.A., M.S.)	More than 20 years	More than 20 years	20	N/A	N/A	
General Education Teacher	School	Female	White	Master's degree (e.g., M.A., M.S.)	11 to 15 years	1 to 5 years	10	Small	Rural	
Special Education Teacher	School	Female	White	Master's degree (e.g., M.A., M.S.)	11 to 15 years	1 to 5 years	9	Large	Urban	
General Education Teacher	School, District	Female	White	Master's degree (e.g., M.A., M.S.)	6 to 10 years	None	7	Medium	Urban	
General Education Teacher	School	Female	White	Master's degree (e.g., M.A., M.S.)	More than 20 years	None	5	Small	Rural	
Parent	Other	Female	White	Bachelor's degree (e.g., B.A., B.S.)	None	None	5	Medium	Rural	Yes
General Education Teacher	School	Female	White	Bachelor's degree (e.g., B.A., B.S.)	11 to 15 years	None	10	Medium	Rural	
General Education Teacher	School	Female	White	Master's degree (e.g., M.A., M.S.)	11 to 15 years	None	3	Medium	Rural	
General Education Teacher	School	Female	White	Bachelor's degree (e.g., B.A., B.S.)	6 to 10 years	1 to 5 years	3	Medium	Rural	

Position	Location of Current Position	Gender	Race/ Ethnicity	Highest Level of Education	Years Teaching Experience	Years Professional Experience	Years Teaching/ Implementing the Montana Science Content Standards	School District Size	School District Urbanicity	Table Leader
General Education Teacher	School	Female	White	Master's degree (e.g., M.A., M.S.)	11 to 15 years	6 to 10 years	12	Medium	Rural	
Special Education Teacher, Co-Teacher in Biology, English, & Pre-Algebra	School	Female	White	Master's degree (e.g., M.A., M.S.)	11 to 15 years	6 to 10 years	11	Large	Urban	
Teacher Naturalist at the Montana Audubon Center	Other	Female	White	Bachelor's degree (e.g., B.A., B.S.)	1 to 5 years	1 to 5 years	9	N/A	N/A	
General Education Teacher	School	Female	White	Master's degree (e.g., M.A., M.S.)	More than 20 years	None	10	Medium	Suburban	

Table A-2. Standard-Setting Panelists, Science Grade 8

Position	Location of Current Position	Gender	Race/ Ethnicity	Highest Level of Education	Years Teaching Experience	Years Professional Experience	Years Teaching/ Implementing the Montana Science Content Standards	School District Size	School District Urbanicity	Table Leader
General Education Teacher	School	Male	White	Master's degree (e.g., M.A., M.S.)	11 to 15 years	1 to 5 years	5	Small	Rural	Yes
Instructional Coach	School	Female	White	Master's degree (e.g., M.A., M.S.)	6 to 10 years	1 to 5 years	4	Large	Urban	
General Education Teacher, Administrator, Parent	School, District	Female	Black or African American	Master's degree (e.g., M.A., M.S.)	16 to 20 years	6 to 10 years	3	Small	Rural	
General Education Teacher	School, District	Female	White	Bachelor's degree (e.g., B.A., B.S.)	1 to 5 years	None	4.5	Large	Urban	
General Education Teacher	School	Female	White	Bachelor's degree (e.g., B.A., B.S.)	More than 20 years	None	9	Medium	Suburban	
General Education Teacher	School, District	Female	White	Master's degree (e.g., M.A., M.S.)	1 to 5 years	None	5	Large	Suburban	
General Education Teacher, Specialist	School	Female	White	Master's degree (e.g., M.A., M.S.)	1 to 5 years	1 to 5 years	3	Large	Urban	
General Education Teacher	School, District	Female	White	Master's degree (e.g., M.A., M.S.)	11 to 15 years	1 to 5 years	7	Small	Rural	
General Education Teacher, Supervising Teacher	School, District	Female	White	Bachelor's degree (e.g., B.A., B.S.)	1 to 5 years	1 to 5 years	5	Small	Rural	

Position	Location of Current Position	Gender	Race/ Ethnicity	Highest Level of Education	Years Teaching Experience	Years Professional Experience	Years Teaching/ Implementing the Montana Science Content Standards	School District Size	School District Urbanicity	Table Leader
Specialist	School	Female	White	Master's degree (e.g., M.A., M.S.)	11 to 15 years	16 to 20 years	10	Small	Rural	Yes
General Education Teacher	School, District	Female	White	Bachelor's degree (e.g., B.A., B.S.)	More than 20 years	None	7	Small	Rural	
General Education Teacher	School	Male	White	Bachelor's degree (e.g., B.A., B.S.)	1 to 5 years	1 to 5 years	1	Large	Urban	
Administrator	Other	Female	White	Master's degree (e.g., M.A., M.S.)	6 to 10 years	6 to 10 years	3	Medium	Suburban	
General Education Teacher	School	Female	White	Master's degree (e.g., M.A., M.S.)	6 to 10 years	None	0	Small	Rural	
General Education Teacher, Mentor	School, District	Female	White	Master's degree (e.g., M.A., M.S.)	More than 20 years	None	16	Large	Suburban	
General Education Teacher	School	Female	White	Master's degree (e.g., M.A., M.S.)	16 to 20 years	None	5	Medium	Rural	

Appendix 3-B

Development of Science Range Performance-Level Descriptors

Development of Science Range Performance-Level Descriptors

1. Development of Three-Dimensional Science Range Performance-Level Descriptors

Cambium Assessment, Inc. (CAI) held a meeting on May 18–19, 2018 for the three-dimensional science assessments. Prior to the meeting, CAI and several client states worked together to refine drafts of Policy and Range PLDs created by Washington State's Office of the Superintendent of Public Instruction (OSPI). During the meeting, educators reviewed and provided feedback on these Policy and Range PLDs.

PLDs describe levels or categories of achievement on a large-scale assessment. PLDs are used to inform the evidence required for item development, inform items selected during the form construction process, and support standard-setting panelist recommendations during the standard-setting process. PLDs are then ultimately used to inform stakeholder interpretation of student scores once standards are set. Egan, Schneider, and Ferrara (2012) recommended four stages of PLD development for the following types of PLDs: Policy, Range, Threshold, and Reporting. The focus of the three-dimensional science PLD meeting was on Policy and Range PLDs only.

2. Definitions of Performance-Level Descriptors by Purpose and Intended Audience

2.1 Policy Performance-Level Descriptors

Policy PLDs articulate the overall claims about a student's achievement in each performance level. They are used by policymakers to broadly articulate the goals and rigor for the state's performance standards. Table 1 shows a sample Policy-based PLD.

Table 1. Draft Three-Dimensional Science Policy PLD for Proficient

Level 3
The Level 3 student is proficient in applying three-dimensional science knowledge and skills as specified in the Next Generation Science Standards (NGSS). The student generally performs at the standard for the grade level, is able to access grade-level content, and engages in higher-order thinking skills with some independence and minimal support.

2.2 Range Performance-Level Descriptors

Range PLDs describe the expectations for students across each Disciplinary Core Idea (DCI) and proficiency level, demonstrating how the content represents a progression of knowledge, skills, and processes across performance levels and grade bands. Washington State's Office of the Superintendent of Public Instruction (OSPI) created Range PLDs for Levels 2, 3, and 4, with Level 3 describing Proficiency. Table 2 shows sample Policy PLDs.

Table 2. Draft NGSS Policy PLDs for Grade 8

Level 2	Level 3	Level 4
Use a model and patterns in data to show that the number of tiny particles does not change during chemical reactions and that particle motion changes when thermal energy is added to or removed from a system.	Develop and use models and interpret patterns in data to show that mass is conserved during chemical reactions and to predict changes in particle motion when thermal energy is added to or removed from a system.	Analyze and interpret patterns in data in order to evaluate and revise a model that describes how mass is conserved during chemical reactions and to explain predicted changes in particle motion when thermal energy is added to or removed from a system.

3. Performance-Level Descriptor Workshop

CAI revised OSPI's PLDs to ensure that text sufficiently differentiates between levels. CAI sent for participating states' review and then convened a committee preparation meeting on May 9, 2018, to prepare participating educators and state staff for the May 18–19th, 2018, meeting.

The meeting was divided into three grade-band rooms: elementary, middle, and high school. One CAI facilitator led each grade-band room, and several CAI staff were available to float between rooms to ensure process consistency and answer questions. Each grade-band room included nine educators, enabling room facilitators to divide the rooms into subgroups to complete the work. Table 3 summarizes the composition of facilitators and educators assigned to each grade band. Recruitment included educators representing special populations (English learners [ELs], Special Education).

Table 3. Workshop Panel Assignments

	Elementary School	Middle School	High School
CAI Facilitators	1	1	1
Educators	9	9	9

3.1 Performance-Level Descriptor Workshop

The Performance-Level Descriptor (PLD) workshop occurred over a period of two days. Appendix 1. PLD Workshop Agenda provides the full workshop agenda.

3.1.1 Day One

The workshop began with a welcome from staff from CAI and participating state staff. CAI provided an overview of the policy aspects of the workshop, including how PLD development fits into the overall test development and standard-setting processes. CAI staff provided training on the processes to be used during the workshop. Following the initial overview, CAI provided training on item clusters and scoring assertions. CAI then described the purpose and structure of the Next Generation Science Standards (NGSS) clusters and scoring assertions, and their importance to the standard-setting process.

A facilitator continued training on Policy PLDs. Facilitators walked panelists through several National Reference Point Policy PLDs, outlining the differences in the key descriptors at each performance level. The panelists reviewed the Policy PLDs individually and in small groups. The panelists used the following questions to frame their review of the National Reference PLDs:

- What terms are used to define proficiency?
- Are there certain terms you value over others?
- Are there words or phrases you note that could inform three-dimensional science policy statements going forward?

After small group discussion, facilitators engaged panelists in a room-level discussion and recorded recommendations for Policy PLDs. Facilitators framed discussions by using the following guiding questions:

- What claims should the Policy PLDs make about students at each performance level?
    - Two to five words that provide context for the expectations of students in each performance level
- What general descriptors best articulate the intended rigor for NGSS?
- How should we represent what proficiency means?
    - College and career readiness
    - On grade-level attainment
    - Meeting standards

The goal of the discussion process is to draft Policy PLDs and for the panelists to begin to have a shared sense of the type of student described by each proficiency level. The Policy PLD discussion lasted through the morning of Day 1, ending with lunch.

After lunch, the meeting shifted to Range PLD training within each breakout room. Facilitators described the process for reviewing Range PLDs. Facilitators modeled how to parse out each PLD, focusing on the key words used in each performance level. In modeling how to parse the standards, the facilitator noted the importance of the Level 3 (proficiency) cut score as an anchor for the other descriptors. The facilitator started by parsing a Level 3, then moving to Levels 2 and 4, modeling the sequence panelists would use throughout the workshop. Next, the facilitator led the room through reviewing one Range PLD. They started by reviewing the Level 3 PLD, then moving to Level 2, then Level 4. Depending on how well the panelists understood the task, the facilitator might have reviewed another PLD with the entire group.

Once the facilitators modeled the process for panelists, they split panelists into groups to create Range PLDs. Each room facilitator divided the PLDs among the groups so they could review them more efficiently in the time allotted for the meeting, resulting in three groups of three panelists in each room. Each group tracked any recommended revisions to each PLD. To facilitate discussion, panelists responded to four questions for each PLD:

- Does the PLD reflect the expected achievement exhibited by students at this performance level?
- What revisions were made to the PLD?
- What rationale do you have for any changes?
- What would distinguish an assertion belonging to this PLD from an assertion belonging to the level below?

One member of each group acted as a scribe, using a computer to track changes to the PLDs, and responded to the questions through an online form. CAI created a template for panelists to use when reviewing the Range PLDs.

For the rest of the afternoon, the panelists reviewed the Range PLDs using the following processes:

- The panelists worked through each assigned PLD, ensuring that the PLD showed a clear progression of observable evidence that should be expected from students at each performance level.
- For each PLD, participants began with the Level 3 descriptor, then moved to Level 2, then Level 4.
- Facilitators monitored progress and work to ensure cross-grade coherence and adherence to the expectations set by the Policy PLDs.

This work continued for the duration of Day 1. At the end of Day 1, CAI and state staff reviewed the panelists' work to check for coherence and consistency across grades.

3.1.2 Day 2

Based on results of the review at the end of Day 1, room facilitators and state staff spent time recalibrating groups if necessary. During the morning of Day 2, the panelists completed their assigned standards. Once each group completed its work, the facilitators conducted discussions with their rooms to ensure coherence across PLDs within each grade band. Each group reviewed their grade-band PLDs to ensure consistency and coherence across performance levels and consistency and coherence within each performance level. This discussion extended until lunch.

After lunch, the grade-band groups met for a cross-grade articulation discussion. They compared the expectations across grade bands to ensure a sensible progression of rigor. The committee focused primarily on examining Level 3 to assess if this level is considered the entry point for college-readiness. After the cross-grade articulation discussion, educators were allowed to adjourn.

For the rest of the afternoon, CAI met with participating state staff. The group discussed the results of the meeting and addressed any issues or inconsistencies in the educators' work. The group also discussed next steps for finalizing the PLDs.

4. References

Egan, K.L., Schneider, M.C., & Ferrara, S. (2012). Performance level descriptors: History, practice, and a proposed framework. In G. Cizek (Ed.), *Setting Performance Standards, Second Edition* (79–106). New York, NY: Routledge.

Science Assessment Team, Office of Superintendent of Public Instruction (2018). *Performance Level Descriptors: Washington Comprehensive Assessment of Science.* Office of Superintendent of Public Instruction.

Schneider, M.C. & Egan, K.L. A Handbook for Creating Range and Target Performance Level Descriptors. The National Center for the Improvement of Educational Assessment.

APPENDIX 1. PLD WORKSHOP AGENDA

Exhibit 1-A. Day 1 PLD Workshop Agenda

Time	Topic	Lead
7:30–8:30 a.m.	Breakfast	
8:30–9:00 a.m.	Welcome Three-Dimensional Science Item Clusters and Scoring Assertions • The purpose and structure of three-dimensional science item clusters • Scoring Assertions	Jon
9:00–9:30 a.m.	Three-Dimensional Science Performance-Level Descriptors (PLDs) • Describe purposes and uses for Policy and Range PLDs • Describe workshop process	Kevin
9:30–9:45 a.m.	Break	
9:45 a.m.–Noon	Policy PLD Discussion • Review Policy PLDs ○ What are the important elements of the descriptor at each performance level? • Small group discussion • Room discussion • Final recommendations	Kevin
Noon–1:00 p.m.	Lunch	

1:00–2:00 p.m.	Range PLD Training
	• Purpose of Range PLDs
	• Tools used in review
	  ○ NGSS Standards
	  ○ Policy PLDs
	  ○ Draft Range PLDs
	  ○ Template for reviewing standards
	• Parsing standards and draft PLDs to differentiate among performance levels
		Room facilitators

2:00–4:30 p.m.	Review draft Range PLDs • Each group reviews assigned PLDs • For each PLD, start with Level 3 (Proficient), then move to Level 2, then Level 4	Room facilitators

Exhibit 1-B. Day 2 PLD Workshop Agenda

Time	Topic	Lead
7:30–8:30 a.m.	Breakfast	
8:30–10:00 a.m.	Continue Range PLD review • Each group reviews assigned PLDs • For each PLD, start with Level 3 (Proficient), then move to Level 2, then Level 4	Room facilitators

10:00 a.m.–Noon	Room Discussion
	• Room discussion to ensure coherence within the grade-band
	  ○ Ensure consistency and coherence across performance levels throughout the grade band
	  ○ Ensure consistency and coherence within each performance level throughout the grade band
		Room facilitators

Noon–1:00 p.m.	Lunch	
1:00–2:30 p.m.	Large group: Cross-grade coherence discussion • Ensure cross-grade consistency and coherence across performance levels • Ensure cross-grade consistency and coherence within each performance level	Kevin
2:30 p.m.	Educators adjourn	
2:30–3:00 p.m.	CAI and Department staff • Resolve inconsistencies within or across grades • Discuss next steps	

Appendix 3-C

Montana Science Assessment Range Performance-Level Descriptors

Montana Science Assessment Range Performance-Level Descriptors

Exhibit C-1. Montana Science Assessment Range Performance-Level Descriptors, Grade 5

Students that are a ____ may be able to do things like...	**Level 1**	**Level 2**	**Level 3**	**Level 4**
Earth and Space Sciences				
ESS1: Earth's Place in the Solar System	Identify data, either in graphical displays or in a model, that would help explain observable features of Earth's landscape, the appearance of stars in the night sky or the patterns created from the orbit and rotation of the Sun-Earth-Moon system and describe how astronomical knowledge is used by American Indians.	Represent data in graphical displays or models, and explain the ordered observable features of Earth's landscape, the appearance of stars in the night sky or the patterns created from the orbit and rotation of the Sun-Earth-Moon system and describe how astronomical knowledge is used by American Indians.	Analyze and interpret graphical displays of data or models to use as evidence in order to explain the ordered, observable features of Earth's landscape, the appearance of stars in the night sky or the patterns created from the orbit and rotation of the Sun-Earth-Moon system and describe how astronomical knowledge is used by American Indians.	Evaluate and revise graphical displays of data or models to make a prediction regarding the ordered, observable features of Earth's landscape, the appearance of stars in the night sky or the patterns created from the orbit and rotation of the Sun-Earth-Moon system and describe how astronomical knowledge is used by American Indians.
ESS2: Earth's Systems	Make observations from data and/or collect information to identify parts of a model and reveal patterns that would show how the interactions between Earth's four major systems might cause patterned features of the Earth, including climate, distribution of water, and physical and biological	Obtain and represent data sets or graphs, and/or carry out investigations using models or information that shows how the interactions between Earth's four major systems might cause patterned features of the Earth, including climate, distribution of water, and physical and biological	Develop and/or use simple models, carry out investigations or evaluate evidence using mathematical thinking, reasoning and information regarding how the interactions between Earth's four major systems might cause patterned features of the Earth, including climate, distribution of water, and physical and biological	Revise a model, analyze the data sets from an investigation using mathematical thinking and research how to better communicate or predict how the interactions between Earth's four major systems might cause patterned features of the Earth, including climate, distribution of water, and

Students that are a ____ may be able to do things like...	Level 1	Level 2	Level 3	Level 4
	constructive and deconstructive forces.	constructive and deconstructive forces.	constructive and deconstructive forces.	physical and biological constructive and deconstructive forces.
ESS3: Earth and Human Activity	Use information and observations from sources to identify either weather-related hazards on humans or human activity on the Earth's resources and environments and describe examples of how American Indians use scientific knowledge and practices to maintain relationships with the natural world.	Identify reliable sources and use obtained information to compare multiple solutions to help explain the cause-and-effect relationship of either weather-related hazards on humans or human activity on the Earth's resources and environments and describe examples of how American Indians use scientific knowledge and practices to maintain relationships with the natural world.	Obtain and use evidence from reliable sources to generate and evaluate the merits or accuracy of a solution that could explain and reduce the cause-and-effect relationship of either weather-related hazards on humans or human activity on the Earth's resources and environments and describe examples of how American Indians use scientific knowledge and practices to maintain relationships with the natural world.	Evaluate, compare and revise a solution to a problem, using evidence obtained from reliable sources, to predict changes that can occur in the cause-and-effect relationships of either weather-related hazards on humans or human activity on the Earth's resources and environments and describe examples of how American Indians use scientific knowledge and practices to maintain relationships with the natural world.
Life Sciences				
LS1: From Molecules to Organisms: Structure and Processes	Identify components of a model that represent parts of a life cycle or behavioral system of organisms; and make observations about organisms that need food for the energy and materials to grow and repair their internal and external structures.	Use a model to that represents the life cycles or behavioral systems of organisms to support an argument; and identify data as evidence to support that organisms need food for the energy and materials to grow and repair their internal and external structures.	Develop and/or use a model to describe patterns in the life cycles or behavioral systems of organisms; and use evidence to construct an argument that organisms need food for the energy and materials to grow and repair their internal and external structures.	Evaluate and revise a model that describes patterns in the life cycles or behavioral systems of organisms when a variable changes; and compare and refine arguments that organisms need food for the energy and materials to grow and repair their

Students that are a ____ may be able to do things like...	**Level 1**	**Level 2**	**Level 3**	**Level 4**
				internal and external structures.
LS2: Ecosystems: Interactions, Energy, and Dynamics	Identify the parts of a model that represents interactions of organisms within an ecosystem and the cycling of matter through those interactions; and identify data that can show how an ecosystem changed.	Use a model to describe the interactions of organisms within an ecosystem and the cycling of matter through those interactions; and collect evidence that shows how an ecosystem can change.	Develop and/or use a model to describe the interactions of organisms within an ecosystem and the cycling of matter through those interactions; and use evidence to show the effect that occurs when one part of the ecosystem is changed.	Evaluate and revise a model that describes the interactions of organisms within an ecosystem and the cycling of matter through those interactions when more information is given; and predict the effects of an ecosystem when one part of the ecosystem is changed.
LS3: Heredity: Inheritance and Variation of Traits	Collect and record data from pictures, drawings, and/or text to help explain that organisms inherit the information that dictates how they look and function; and make an observation about an organism when its environment changes.	Use data collected from tables and various graphical displays to support an explanation that organisms inherit the information that dictates how they look and function; and identify information that would help explain what happens to an organism if the environment changes.	Analyze and interpret various forms of data to construct an explanation that organisms inherit the information that dictates how they look and function; and construct an explanation using evidence that supports that an organism has changed in response to environmental changes.	Construct, analyze and interpret tables and graphical displays of data in order to construct and revise an explanation that organisms inherit the information that dictates how they look and function; and predict what would happen to an organism if its environment continues to change.

Students that are a ____ may be able to do things like…	**Level 1**	**Level 2**	**Level 3**	**Level 4**
LS4: Biological Evolution: Unity and Diversity	Identify patterns in past or present organism characteristics that can be used as evidence to support that when there is a change in the environment, certain individual organisms could have variations in traits that lead to advantages in survival and reproduction; and use observations from pictures, drawings, and/or writings to support that current, living organisms can only survive in particular environments or resemble organisms that once lived on Earth.	Identify and/or record past and present observations that could either provide evidence that when there is a change in the environment, certain individual organisms could have variations in traits that lead to advantages in survival and reproduction, or that living organisms resemble organisms that once lived on earth; and identify data that can be used to compare the merits of a solution that can affect a population of organisms.	Analyze and interpret past and present organism characteristics to either provide evidence that when there is a change in the environment, certain individual organisms could have variations in traits that lead to advantages in survival and reproduction, or that living organisms resemble organisms that once lived on earth; and analyze and compare the merits of a solution that can affect a population of organisms.	Analyze and interpret past and present organism characteristics to evaluate and revise a constructed explanation that states that with a change in the environment, certain individual organisms could have variations in traits that lead to advantages in survival and reproduction, or that living organisms resemble organisms that once lived on earth; and compare sets of data to help argue the merits of a solution that could affect a population of organisms.
Physical Sciences				
PS1: Matter and Its Interactions	Make observations about variables that are controlled to determine if a chemical reaction occurs and a new substance is created, measuring and graphing quantities to show matter is always conserved regardless of the change that occurs; and to develop a model to show matter exists made of particles too small to be seen.	Use models testing variables that are controlled to determine if a chemical reaction occurs and a new substance is created, measuring and graphing quantities to show matter is always conserved regardless of the change that occurs; and to develop a model to show matter exists made of particles too small to be seen.	Conduct an investigation in which variables are controlled to determine if a chemical reaction occurs and a new substance is created, measuring and graphing quantities to show matter is always conserved regardless of the change that occurs; and to develop a model to show matter exists made of particles too small to be seen.	Evaluate and revise a model using quantitative data in which variables are controlled to determine if a chemical reaction occurs and a new substance is created, measuring and graphing quantities to show matter is always conserved regardless of the change that occurs; and to develop a model to show matter

Students that are a ____ may be able to do things like...	Level 1	Level 2	Level 3	Level 4
				exists made of particles too small to be seen.
PS2: Motion and Stability: Forces and Interactions	Identify questions from an investigation about cause-and-effect relationships between balanced and unbalanced forces (magnetism and/or gravity) and an object's motion.	Use observations from an investigation to provide evidence creating an argument about cause-and-effect relationships between balanced and unbalanced forces (magnetism and/or gravity) and an object's motion.	Ask questions, plan and conduct an investigation and/or use produced data to provide evidence supporting an argument about cause-and-effect relationships between balanced and unbalanced forces (magnetism and/or gravity) and an object's motion.	Ask questions, conduct and compare two different investigations and/or use produced data to provide evidence to predict cause and effect relationships between balanced and unbalanced forces (magnetism and/or gravity) and an object's motion.
PS3: Energy	Make observations using produced data to ask questions on how energy can be used as a fuel or food; or be transferred from stored and/or motion energy to different forms like sound, light, and electrical currents.	Make observations using produced data to provide evidence on how energy can be used as a fuel or food; or be transferred from stored and/or motion energy to different forms like sound, light, and electrical currents.	Use models, ask questions to make predictions and/or use produced data to provide evidence on how energy can be used as a fuel or food; or be transferred from stored and/or motion energy to different forms like sound, light, and electrical currents.	Evaluate and revise models, ask questions to make predictions and/or use produced data to provide evidence on how energy can be used as a fuel or food; or be transferred from stored and/or motion energy to different forms like sound, light, and electrical currents.
PS4: Waves and their Applications in Technologies for Information Transfer	Make observations about patterns of light or mechanical waves using models; and use evidence to explain how reflected light from objects causes objects to be seen. Construct and compare multiple solutions to transfer information.	Develop and/or use a model to describe the patterns of light or mechanical waves; and use evidence to explain how reflected light from objects causes objects to be seen. Construct and compare multiple solutions to transfer information.	Create a solution or develop and/or use a model to describe the patterns of light or mechanical waves; and use evidence to explain how reflected light from objects causes objects to be seen. Construct and compare	Revise a model to make predictions and describe the patterns of light or mechanical waves; and use evidence to explain how reflected light from objects causes objects to be seen. Construct and compare

Students that are a ____ may be able to do things like...	Level 1	Level 2	Level 3	Level 4
			multiple solutions to transfer information.	multiple solutions to transfer information.

Exhibit C-2. Montana Science Assessment Range Performance-Level Descriptors, Grade 8

Students that are a _____ may be able to do things like…	**Level 1**	**Level 2**	**Level 3**	**Level 4**
Earth and Space Sciences				
ESS1: Earth's Place in the Solar System	Identify components of a model that measures and collects evidence to explain the similarities and differences in the patterned motions of the Sun-Earth-Moon system, the role of gravity in the motion of galaxies and the solar system, or the relative occurrence of events in the Earth’s and solar system’s history.	Use a model to identify data from tables and other graphical displays that can be used as pieces of evidence to explain the patterned motions of the Sun-Earth-Moon system, the role of gravity in the motion of galaxies and the solar system, or the relative occurrences of events in the Earth’s and solar system’s history.	Develop and/or use a model using graphical displays of data that explain the patterned motions of the Sun-Earth-Moon system, the role of gravity in the motion of galaxies and the solar system, or the relative occurrence of events in the Earth’s and solar system’s history.	Evaluate and revise a model based on constraints and data limitations that explain the patterned motions of the Sun-Earth-Moon system, the role of gravity in the motion of galaxies and the solar system, or the relative occurrence of events in the Earth’s and solar system’s history.
ESS2: Earth's Systems	Take measurements and/or make observations from graphical data to help identify the components of a model that help explain the patterns in the flow or cycles of energy and matter throughout Earth’s systems, including the sun and Earth’s interior as primary energy sources; and identify evidence to explain that Earth’s processes have changed the Earth’s surface at varying spatial and time scales.	Use a model or investigation to identify patterns from bar graphs, pictographs, and other various graphical data that supports how energy and matter flow or cycle throughout Earth’s systems, including the sun and Earth’s interior as primary energy sources; and organize evidence to explain how Earth’s processes have changed the earth’s surface at varying spatial and time scales.	Analyze data from an investigation to develop, use and/or revise a model that shows patterns in the flow or cycles of energy and matter throughout Earth’s systems, including the sun and Earth’s interior as primary energy sources; and interpret evidence to construct an explanation for how Earth’s processes have changed the Earth’s surface at varying spatial and time scales.	Evaluate and revise a model to generate data that supports an explanation that shows patterns in how energy and matter flow or cycle throughout Earth’s systems, including the sun and Earth’s interior as primary energy sources; and evaluate the impact of new data by predicting how the Earth’s processes will change the earth’s surface at varying spatial and time scales if a new variable is introduced.

Students that are a _____ may be able to do things like…	**Level 1**	**Level 2**	**Level 3**	**Level 4**
ESS3: Earth and Human Activity	Identify scientific questions using collected and/or graphically represented evidence regarding the dependency of humans on the environment for different resources; and identify evidence that can help design a simple solution that minimizes the effect of humans on the environment or identify the observed patterns that emerge between natural hazards and their related geological forces.	Ask questions about data or apply scientific ideas, about the uneven distribution of natural resources and human dependence on the environment for those resources, to design a simple solution that minimizes the effect of humans on the environment; and to explain the observable patterns seen in the data from the history of natural hazards and their related geological forces.	Analyze and interpret sets of data regarding the uneven distribution of natural resources and human dependence on the environment for those resources to ask questions and design a solution that could minimize the effect of humans on the environment; and to explain the observable patterns seen in the data from the history of natural hazards and their related geological forces.	Analyze and interpret sets of data regarding the uneven distribution of natural resources and human dependence on the environment for those resources to evaluate and revise a question that can modify a design solution that minimizes the effect of humans on the environment; and explains the effect of humans on the environment; and predicts future patterns of natural hazards when considering the impact of humans on the environment.
Life Sciences				
LS1: From Molecules to Organisms: Structure and Processes	Collect data from an investigation to support an argument using evidence or identify the components of a model that explains that all living things are made up of cells that work together to form more complex structures and systems; both plants and animals convert energy into food sources but the process to do so is different;	Collect and organize data from an investigation to support an argument using evidence or use a model to explain that all living things are made up of cells that work together to form more complex structures and systems; both plants and animals convert energy into food sources but the process to do so is different;	Collect and synthetize data from an investigation to engage in an argument using evidence or develop/use a model to explain that all living things are made up of cells that work together to form more complex structures and systems; both plants and animals convert energy into food sources but the process to do so is different;	Evaluate and revise a model or explanation using investigative data as evidence to engage in an argument that all living things are made up of cells that work together to form more complex structures and systems; both plants and animals convert energy into food sources but the process to do so is

Students that are a ____ may be able to do things like...	Level 1	Level 2	Level 3	Level 4
	characteristic animal behaviors and specialized plant structures affect the probability of reproduction.	characteristic animal behaviors and specialized plant structures affect the probability of reproduction.	characteristic animal behaviors and specialized plant structures affect the probability of reproduction.	different; characteristic animal behaviors and specialized plant structures affect the probability of reproduction.
LS2: Ecosystems: Interactions, Energy, and Dynamics	Identify components of a model that explains the dynamic relationships and interactions between the diverse types of living and nonliving parts of an ecosystem including the flow of energy and the cycling of matter among organisms and abiotic components of an ecosystems; and organize multiple graphical displays of data to support a solution to mitigate disruptions to any part of an ecosystem by human access to natural resources; and analyze scientific concepts used by American Indians to maintain healthy relationships with environmental sources.	Use a model to explain the dynamic relationships and interactions between the diverse types of living and nonliving parts of an ecosystem including the flow of energy and the cycling of matter among organisms and abiotic components of an ecosystems; and organize and identify patterns from multiple graphical displays of data to support a solution to mitigate disruptions to any part of an ecosystem by human access to natural resources; and analyze scientific concepts used by American Indians to maintain healthy relationships with environmental sources.	Develop and/or use a model to explain and predict the dynamic relationships and interactions between the diverse types of living and nonliving parts of an ecosystem including the flow of energy and the cycling of matter among organisms and abiotic components of an ecosystems; and analyze and interpret multiple graphical displays of data to design and support a solution to mitigate disruptions to any part of an ecosystem by human access to natural resources; and analyze scientific concepts used by American Indians to maintain healthy relationships with environmental sources.	Analyze and revise a model that explains and supports the dynamic relationships and interactions between the diverse types of living and nonliving parts of an ecosystem, including the flow of energy and the cycling of matter among producers, consumers, and decomposers when a variable in the system is changed; and evaluate limitations of data when analyzing and interpreting multiple graphical displays of data to design and support a solution to mitigate disruptions to any part of an ecosystem by human access to natural resources; and analyze scientific concepts used by American Indians to maintain healthy relationships with environmental sources.

Students that are a _____ may be able to do things like…	Level 1	Level 2	Level 3	Level 4
LS3: Heredity: Inheritance and Variation of Traits	Identify the components of a model that describes the relationship among variables that show why sexual/asexual reproduction may have different results of genetic variation in offspring; and how complex and microscopic structural changes to genes (mutations) can be analyzed to determine how they affect the structure and function of an organism.	Use a model to describe the relationship among variables that show why sexual/asexual reproduction may have different results of genetic variation in offspring; and how complex and microscopic structural changes to genes (mutations) can be analyzed to determine how they affect the structure and function of an organism.	Develop and/or use a model to describe the relationship among variables that show why sexual/asexual reproduction may have different results of genetic variation in offspring; and how complex and microscopic structural changes to genes (mutations) can be analyzed to determine how they affect the structure and function of an organism.	Evaluate and revise a model that either describes the relationship among variables as to why sexual/asexual reproduction may have different results of genetic variation in offspring or predicts what changes would occur in the function of an organisms if there is a mutation in the organism's genes.
LS4: Biological Evolution: Unity and Diversity	Identify the patterns in large data sets to explain why species can change over time and communicate the similarities or differences found in past and present organisms or fossil records of past environmental conditions; and use data to construct an explanation for how humans influence the biodiversity of an area, and natural or artificial selection can give some organisms an advantage in survival and reproduction.	Identify and organize the patterns in large data sets to explain why species can change over time and communicate the similarities or differences found in past and present organisms or fossil records of past environmental conditions; and gather and use data to construct an explanation for how humans influence the biodiversity of an area, and natural or artificial selection can give some organisms an advantage in survival and reproduction.	Analyze and interpret the patterns in large data sets to explain why species can change over time and communicate the similarities or differences found in past and present organisms or fossil records of past environmental conditions; and gather and synthesize data to construct an explanation for how humans influence the biodiversity of an area, and natural or artificial selection can give some organisms an advantage in survival and reproduction.	Analyze and evaluate an explanation of the patterns in large data sets that show the similarities or differences found in past and present organisms or fossil records of past environmental conditions; and apply concepts of statistics and probability (variability) to form an explanation that as humans influence the biodiversity of an area, natural or artificial selection can give some organisms an advantage in survival and reproduction.
Physical Sciences				

Students that are a ____ may be able to do things like...	Level 1	Level 2	Level 3	Level 4
PS1: Matter and Its Interactions	Identify the components a model that explains the conservation of mass when two substances react; and collect data to construct an explanation using evidence that supports that the properties of matter are a function of the composition of atoms and molecules that make up matter, as well as the thermal energy.	Use a model to explain the conservation of mass when two substances react; and collect and interpret data to construct an explanation using evidence that supports that the properties of matter are a function of the composition of atoms and molecules that make up matter, as well as the thermal energy.	Develop and/or use a model to explain the conservation of mass when two substances react; and collect, interpret, and analyze data to construct an explanation using evidence that supports that the properties of matter are a function of the composition of atoms and molecules that make up matter, as well as the thermal energy.	Evaluate and revise a model to explain the conservation of mass when two substances react; and use evidence to predict how changes to the molecular structure or thermal energy of matter can affect its properties.
PS2: Motion and Stability: Forces and Interactions	Identify questions, conduct an investigation, and collect data, regarding the relationship between mass, force, and motion, and the attractive and repulsive forces that act at a distance (electric, magnetic, and gravitational forces).	Identify questions, conduct an investigation, and organize and use data to make a claim regarding the relationship between mass, force, and motion, and the attractive and repulsive forces that act at a distance (electric, magnetic, and gravitational forces).	Ask questions, design and conduct an investigation, and analyze and interpret data to make and support a claim regarding the relationship between mass, force, and motion, and the attractive and repulsive forces that act at a distance (electric, magnetic, and gravitational forces).	Ask questions, conduct, evaluate, and revise an investigation, and analyze and evaluate data to make and support a claim regarding the relationship between mass, force, and motion, and the attractive and repulsive forces that act at a distance (electric, magnetic, and gravitational forces).

Students that are a ____ may be able to do things like...	Level 1	Level 2	Level 3	Level 4
PS3: Energy	Identify components of a model that investigates how kinetic and potential energy interact, transform, or transfer to another object; and collect data for an investigation that provides data regarding the temperature and total energy of a system and its dependency on a variety of factors, including the types and states of matter, as well as the amount of matter involved.	Use a model or investigation to describe kinetic and potential energy interact, transform, or transfer to another object; and collect and organize data for an investigation regarding the temperature and total energy of a system and its dependence on a variety of factors, including the types and states of matter, as well as the amount of matter involved.	Develop and/or use a model or investigation to describe how kinetic and potential energy interact, transform, or transfer to another object; and analyze data from an investigation to provide evidence that the temperature and total energy of a system is dependent on a variety of factors, including the types and states of matter, as well as the amount of matter involved.	Evaluate and revise a model or investigation to predict changes to the interaction of kinetic and potential energy, including how energy is transformed, or transferred to another object; and apply concepts of statistics and probability when providing evidence that the temperature and total energy of a system is dependent on a variety of factors, including the types and states of matter, as well as the amount of matter involved.
PS4: Waves and their Applications in Technologies for Information Transfer	Identify the mathematical components in a model that describe the patterns observed between wave characteristics and wave energy; and select a claim with evidence to show that waves are reflected, absorbed, or transmitted through various materials.	Use mathematical representations in a model to describe the patterns observed between wave characteristics and wave energy; and support a claim with evidence to show that waves are reflected, absorbed, or transmitted through various materials.	Develop and/or use mathematical representations in a model to describe the patterns observed between wave characteristics and wave energy; and construct a claim supported by evidence to show that waves are reflected, absorbed, or transmitted through various materials.	Evaluate and revise a mathematical model to predict patterns between wave characteristics and wave energy; and integrate qualitative, quantitative, and technical data to provide evidence to support a claim that waves are reflected, absorbed, or transmitted through various materials.

Appendix 3-D

Standard-Setting Assertion Maps

Standard-Setting Assertion Maps

Exhibit D-1. Standard-Setting Assertion Map, Science Grade 5

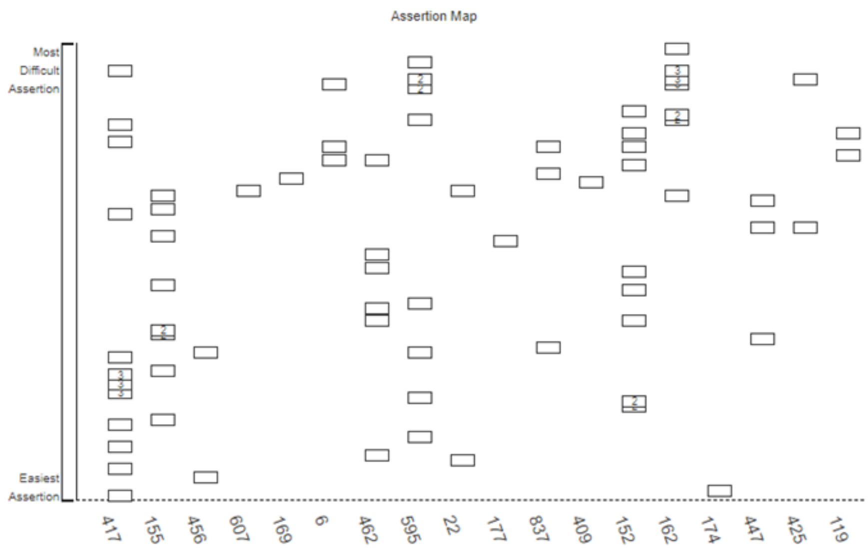
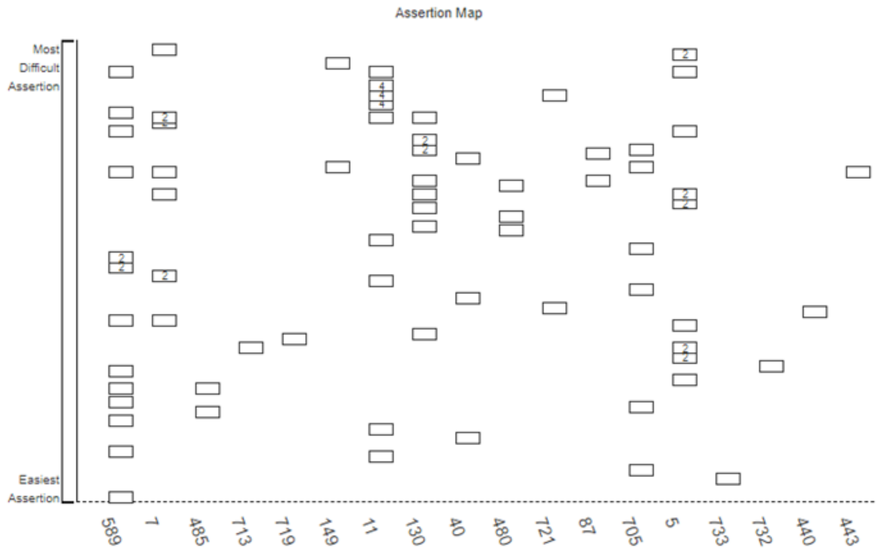


Exhibit D-2. Standard-Setting Assertion Map, Science Grade 8



Appendix 3-E

Standard-Setting Workshop Agenda

Standard-Setting Workshop Agenda

Exhibit E-1. Day 1 Standard-Setting Workshop Agenda

2022 Standard Setting for the Montana Science Assessment

SCIENCE PANEL AGENDA

August 2 – 3, 2022

Standard-Setting Workshop Day 1 – Tuesday, August 2, 2022

8:00 – 8:30 a.m.	**Participant Login**
8:30 – 8:45 a.m.	**Welcome and Introductions from the Montana Office of Public Instruction**

8:45 – 9:30 a.m.	**Large-Group Orientation and Introductory Training**
	Welcome and introductions
	Purpose of standard-setting workshop
	General overview of standard-setting procedures and key concepts
	• Performance-level descriptors (PLDs)
	• Item clusters and stand-alone items
	  Item interactions
	  Scoring assertions
	• Item cluster review
	• Assertion mapping – two rounds
	• Contextual information – benchmark and impact data
	• Panelist feedback and impact data

9:30 – 9:45 a.m.	**Break, and Separate into Small Group Rooms**
9:45 – 11:15 a.m.	**Panelists Experience Online Operational Assessment and Test Environment**
11:15 – 12:15 p.m.	**Review Range PLDs and Discuss Threshold PLDs** Parse range PLDs to identify specific claims within performance levels Identify knowledge and skills differentiating student performance between levels
12:15 – 1:00 p.m.	**Lunch (on your own)**
1:00 – 2:00 p.m.	**Continue Discussions of PLDs**
2:00 – 5:00 p.m.	**Review of Ordered Scoring Assertion Booklet (OSAB) Items** Composition of the item clusters and stand-alone items Training on how to review item clusters and stand-alone items • How do the item interactions support the scoring assertion? • Why is this assertion more difficult than the previous assertion? • How does the scoring assertion and the underlying interactions relate to the PLDs? Training on usage of contextual information – benchmark and impact data Instruction in accessing the item clusters and stand-alone items Review of item clusters and stand-alone items in the OSAB
5:00 p.m.	**Adjourn**

Cambium Assessment, Inc. 1

Exhibit E-2. Day 2 Standard-Setting Workshop Agenda

2022 Standard Setting for the Montana Science Assessment

SCIENCE PANEL AGENDA

August 2 – 3, 2022

Standard-Setting Workshop Day 2 – Wednesday, August 3, 2022	
8:30 – 10:00 a.m.	**Continue Review of OSAB Items**
10:00 – 11:00 a.m.	**Training on Assertion-Mapping Task** Review of assertion-mapping key concepts • Performance-level descriptors (PLDs) • Ordered scoring assertions • Assertion map • Contextual information – benchmark and impact data Training on assertion-mapping tool Practice assertion-mapping task and standard-setting quiz
11:00 – 11:15 a.m.	**Break**
11:15 – 12:30 p.m.	**Round 1 Assertion Mapping** Review of assertion-mapping procedures and key concepts Completion of assertion-mapping readiness form Round 1 assertion mapping
12:30 – 1:30 p.m.	**Lunch (on your own)**
1:30 – 2:30 p.m.	**Review Panelist Feedback Data and Discuss Round 1 Results** How to use panelist agreement feedback data and impact data Presentation and discussion of Round 1 panelist agreement feedback data and impact data
2:30 – 3:30 p.m.	**Round 2 Assertion Mapping** Review of assertion-mapping procedures and key concepts Completion of assertion-mapping readiness form Round 2 assertion mapping
3:30 – 4:00 p.m.	**Workshop Evaluations and Educator Panel Adjourn**
4:00 – 5:00 p.m.	**Across Grade Moderation with All Science Table Leaders**
5:00 p.m.	**Table Leader Adjourn**

Appendix 3-F

Standard-Setting Training Slides

Exhibit F-1. Large-Group Orientation Slides

Standard Setting: Science

August 2 – 3, 2022
Montana Science Assessments (MSA)

2

Welcome and Introductions

Montana Office of Public Instruction (OPI)

State Representatives

3

Montana Office of Public Instruction (OPI)

- Ashley McGrath
- Katie Murnion
- Chris Noel
- Michelle McCarthy
- Jennifer Stadum

4

Large-Group Orientation

Cambium Assessment, Inc.

Workshop Leaders

5

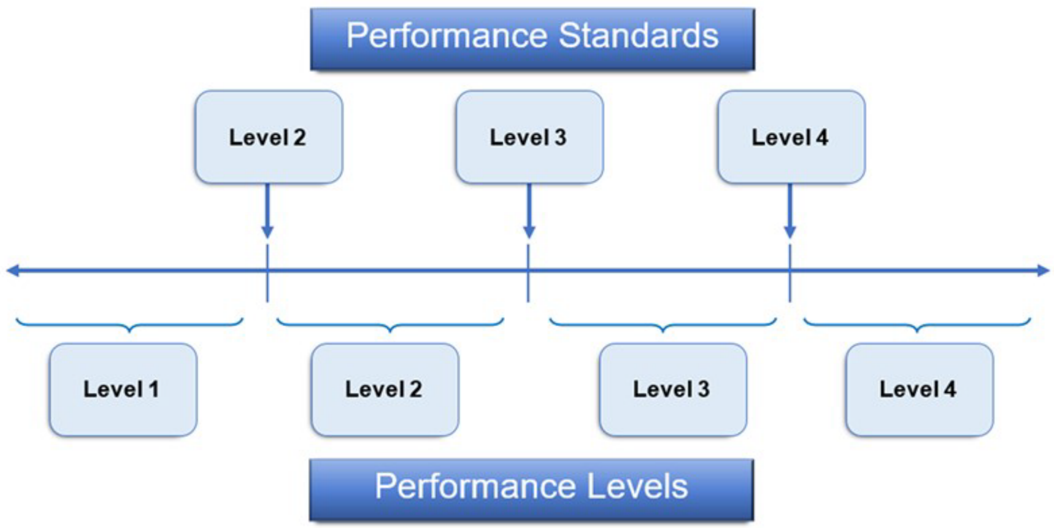
Cambium Assessment, Inc.

- Psychometrics
  - Stephan Ahadi
  - Zhongtian Lin
  - Zebing Wu
- Room Facilitators
  - Grade 5: Hibbah Haddam and Olivia Francois
  - Grade 8: Matthew Davis and Jared Taylor

Purpose of the Standard-Setting Workshop

6

- Recommend to the Montana Board of Public Education (BPE) three performance standards to differentiate the four performance levels on the MSA in grades 5 and 8



Main Workshop Activities

7

- Large-Group Orientation
- Panel Training
  - Take the Online Operational Assessment
  - Review Range PLDs
  - Discuss Just Barely PLDs
  - Review the Ordered Scoring Assertion Booklet
  - Training on Assertion-Mapping Procedure
- Recommend Performance Standards
  - Two rounds
  - Panelist feedback following Round 1
  - Vertical Articulation
- Workshop Evaluation

Importance of Security

8

Please do not:

- Create any form of electronic copy of test content (screenshots, electronic notes, etc.)
- Create any hand-written notes of test content
- Discuss test content with anyone outside the meeting
- Use your computer during the course of the meeting for any purpose other than participating in the item review (e.g., email, web browsing, social media)
- Save notes about item or passage content to your computer

Cambium Assessment

OFFICE OF PUBLIC INSTRUCTION

Reason for New Science Standards

9

- The Montana Board of Public Education adopted the Montana Science Content Standards in September 2016
- In spring 2022, assessments aligned to the 2016 Montana Science Content Standards were administered to all students in Montana in grades 5 and 8

Description of the Science Test Design

10

- Grades 5 and 8 tests assess students' understanding of the Montana Science Content Standards
- The MSA at each grade includes 6 item clusters and 12 stand-alone items
  - **Item clusters** include a stimulus and a series of questions that generally take students about 6–12 minutes to complete
  - **Stand-alone items** are shorter and generally take 1–3 minutes to complete
- All items ask students to use science and engineering practices and apply their understanding of disciplinary core ideas and crosscutting concepts to make sense out of real-world phenomena

Cambium Assessment

Scoring Assertions

11

- Within each item cluster, a series of explicit assertions can be made about the knowledge and skills that a student has demonstrated based on specific features of the student's responses
- Scoring assertions can be supported based on students' responses in one or more interactions within an item cluster.
- For example:
  - A student correctly graphs data points indicating that (s)he can construct a graph showing the relationship between two variables,
  - Makes an incorrect inference about the relationship between the two variables, thereby not supporting the assertion that the student can interpret relationships expressed graphically

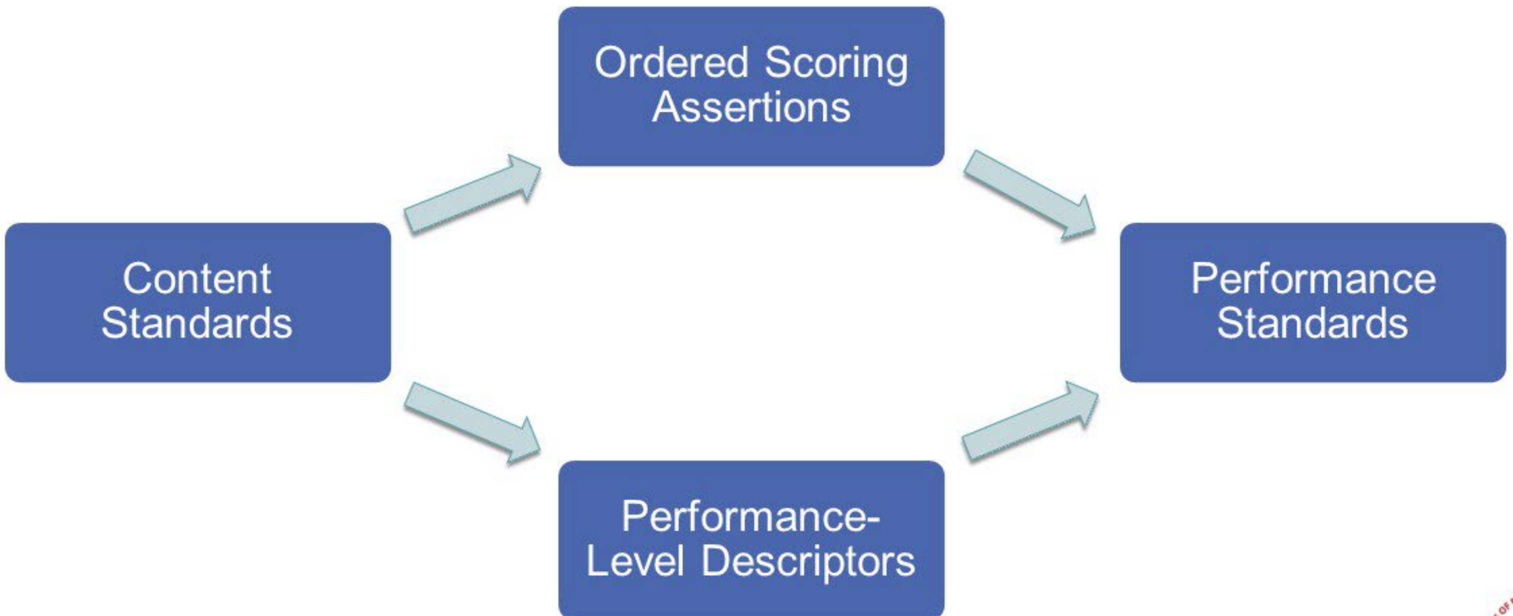
Standard Setting

12

- Systematic process by which trained participants use their knowledge of academic content standards, test items, and student performance to recommend cut-scores associated with each performance level on the test

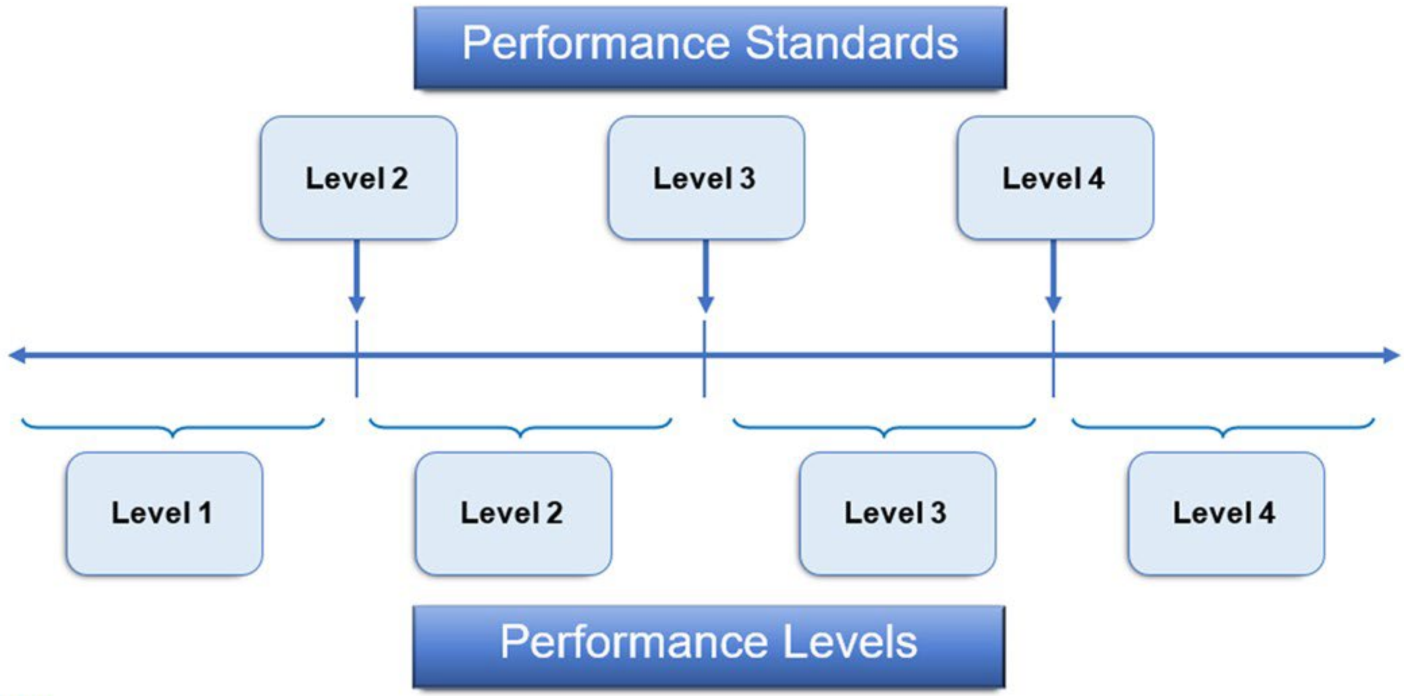
From Content Standards to Performance Standards

13



Performance Standards and Performance Levels

14



Assertion-Mapping Procedure (AMP)

15

- Test-centered procedure
- Employs an ordered item procedure adapted to accommodate new multiple interaction item types
- Map ordered scoring assertions to performance levels
- Is being employed to recommend performance standards in multiple states assessing three-dimensional science standards

Cambium Assessment

OFFICE OF PUBLIC INSTRUCTION

Key Elements of the AMP

16

- Performance-level descriptors (PLDs)
  - Range PLDs
  - Threshold PLDs (just barely meets)
- Ordered scoring assertions
- Assertion map
- Assertion mapping in multiple rounds
  - Contextual information – benchmarking data and student impact data
  - Panelist feedback and group discussion
- Vertical articulation and moderation

Performance-Level Descriptors (PLDs)

17

- Describe what students within each performance level are expected to know and be able to do
- PLDs are the link between the content and performance standards

Grade 8 Range PLDs – Level 3

18

Physical Sciences

- **MS-PS1:** Develop and/or use a model to explain the conservation of mass when two substances react…
- **MS-PS2:** Ask questions, design and conduct an investigation, and analyze and interpret data to make and support a claim regarding the relationship between mass, force, and motion, and the attractive and repulsive forces that act at a distance (electric, magnetic, and gravitational forces.)
- **MS-PS3:** Develop and/or use a model or investigation to describe how kinetic and potential energy interact, transform, or transfer to another object…
- **MS-PS4:** Develop and/or use mathematical representations in a model to describe the patterns observed between wave characteristics and wave energy...

Grade 8 Range PLDs Across Performance Levels

19

MS-PS3: Energy

- **Level 1:** **Identify components of a model** that investigates how kinetic and potential energy interact, transform, or transfer to another object; and **collect data for an investigation** that provides data regarding the temperature and total energy of a system and its dependency on a variety of factors, including the types and states of matter, as well as the amount of matter involved.
- **Level 2:** **Use a model or investigation to describe** kinetic and potential energy interact, transform, or transfer to another object; and **collect and organize data for an investigation** regarding the temperature and total energy of a system and its dependence on a variety of factors, including the types and states of matter, as well as the amount of matter involved.
- **Level 3:** **Develop and/or use a model or investigation** to describe how kinetic and potential energy interact, transform, or transfer to another object; and **analyze data from an investigation** to provide evidence that the temperature and total energy of a system is dependent on a variety of factors, including the types and states of matter, as well as the amount of matter involved.
- **Level 4:** **Evaluate and revise a model or investigation to predict** changes to the interaction of kinetic and potential energy, including how energy is transformed, or transferred to another object; and **apply concepts of statistics and probability when providing evidence** that the temperature and total energy of a system is dependent on a variety of factors, including the types and states of matter, as well as the amount of matter involved.

Cambium Assessment

Important Concepts

20

- "Just barely" meets the performance level
  - Differentiate students who just barely qualify for entry into a performance level from those just below
- Assertion mapping
  - Map each scoring assertion to the performance level that the assertion best supports
- Ordering of assertions
  - Assertions are ordered by difficulty within an item
  - Mapping of assertions to performance levels should reflect the ordering – no inversions within an item

Cambium Assessment

OFFICE OF PUBLIC INSTRUCTION

Ordered Scoring Assertions

21

- The ordered scoring assertion booklet (OSAB) constitutes a test administration:
  - A test form that meets test blueprint specifications
- It is important to evaluate scoring assertions as they relate to the item interactions
- Assertions within items are ordered by difficulty
  - Assertions within an item may not represent all PLDs

What If an Assertion Seems Out of Order?

22

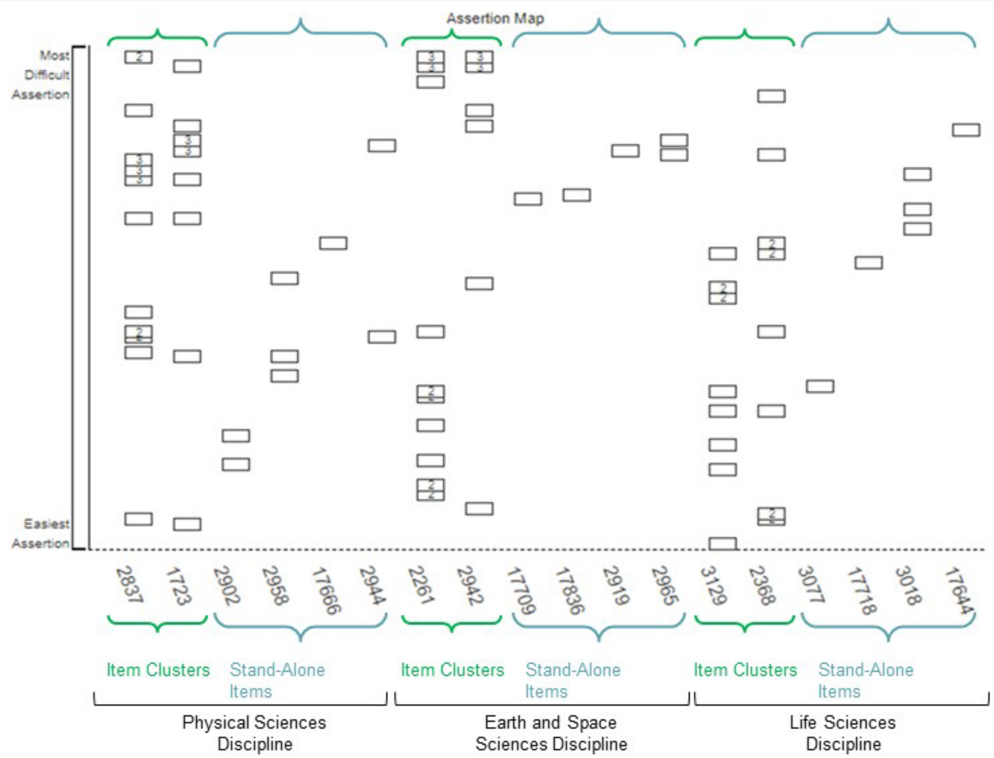
- Assertion ordering is based on student performance
- Assertions may seem out of order because they are ordered by difficulty, and not by content or cognitive process
- Identify why a scoring assertion is more difficult than the assertions before it, and easier than the assertions following it
  - Pay special attention to the interactions supporting the assertions
  - Assertions may be more or less difficult because of the underlying interactions

Cambium Assessment

OFFICE OF PUBLIC INSTRUCTION

Assertion-Mapping Task

23



Studying the Items and Scoring Assertions

24

- Working individually, for each scoring assertion ask yourself:
  1. *How do the item interactions support the scoring assertion?*
  2. *Why is this assertion more difficult than the previous assertions (within the item)?*
  3. *How does the scoring assertion and the underlying interactions relate to the PLDs?*
- Working as a group
  - Discuss how item interactions support scoring assertions
  - Discuss ordering of scoring assertions
  - Discuss how scoring assertions are related to the PLDs

Cambium Assessment

OFFICE OF PUBLIC INSTRUCTION

What If an Item Seems Wrong or Unfair?

25

- Do not let yourself get distracted – this is not an item review meeting
- If you believe something is wrong with an item interaction or scoring assertion, tell the Workshop Leader, then skip over the assertion as you review the rest of the assertions within the item

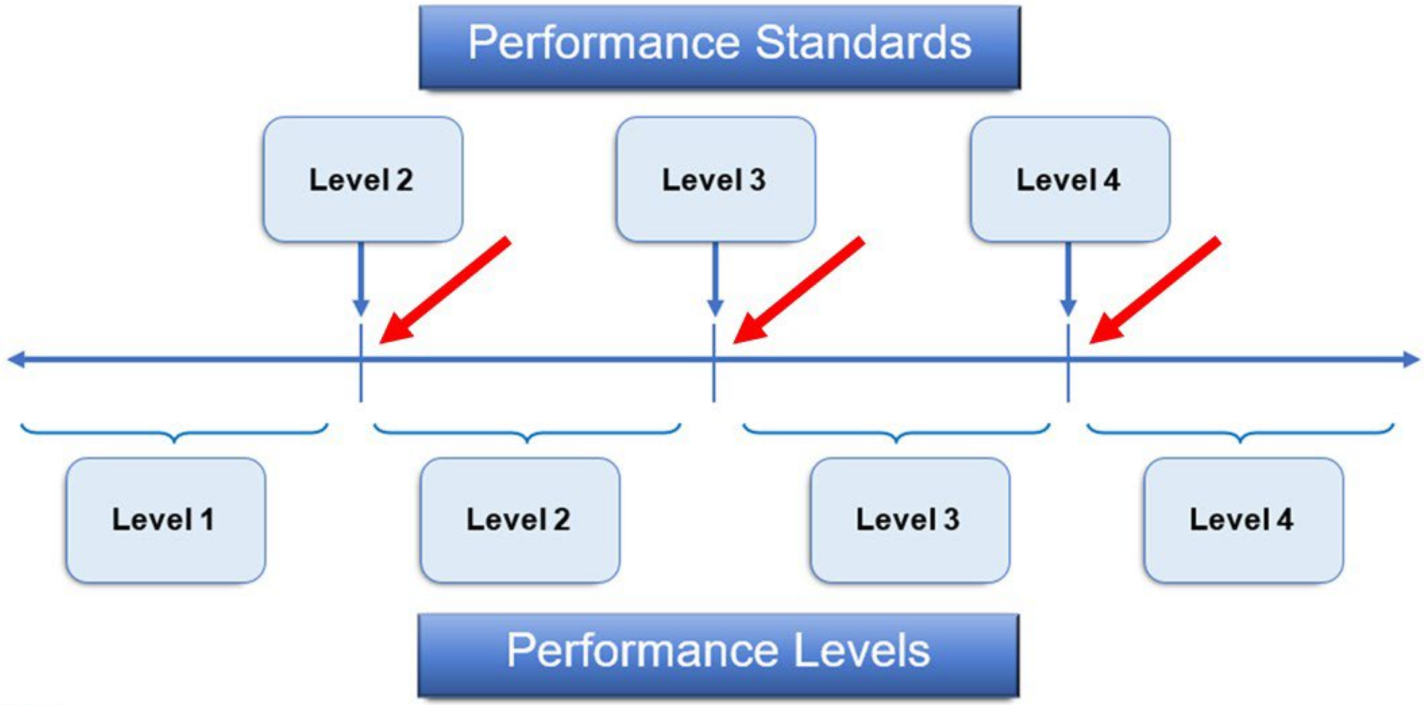
"Just Barely" Meets the Performance Standard

26

- When considering each performance level, we are especially interested in the transition areas between performance levels
- Pay attention to characteristics of students who ***just barely*** qualify for entry into the performance level from those just below
  - Not a typical example of students in the performance level
  - Although they are not good examples of the performance level, they do still meet the standard, or description in the PLD

Threshold “Just Barely” PLDs

27



Assertion-Mapping Task

28

- Map assertions to performance levels
  - Consider what differentiates students who just barely qualify for entry into the performance level from those not quite ready for entry into the performance level
  - Evidence that the student has demonstrated knowledge and skills necessary for entry into the performance level
- Map assertions in the online standard-setting tool

Group Feedback and Discussion

29

- Goals
  - Add important information to your thinking
  - Develop common understandings
  - Inform possible re-evaluation of assertion mappings
- Expectation is converging judgments
  - Consensus is not a requirement or goal

Feedback and Impact Data

30

- Percentage of students reaching or exceeding the standard based on assertion mappings
- Group discussion
  - Does the percentage of students reaching or exceeding the current recommended performance standard seem reasonable?
  - What are the implications for the performance standards?
  - All performance standard recommendations should be based on content rationales drawn from the Montana Science Content Standards

Creating a System of Performance Standards

31

- Performance standards for a statewide system must be coherent across grades and subjects
  - Articulation
  - Benchmarking
  - Moderation

Benchmarking

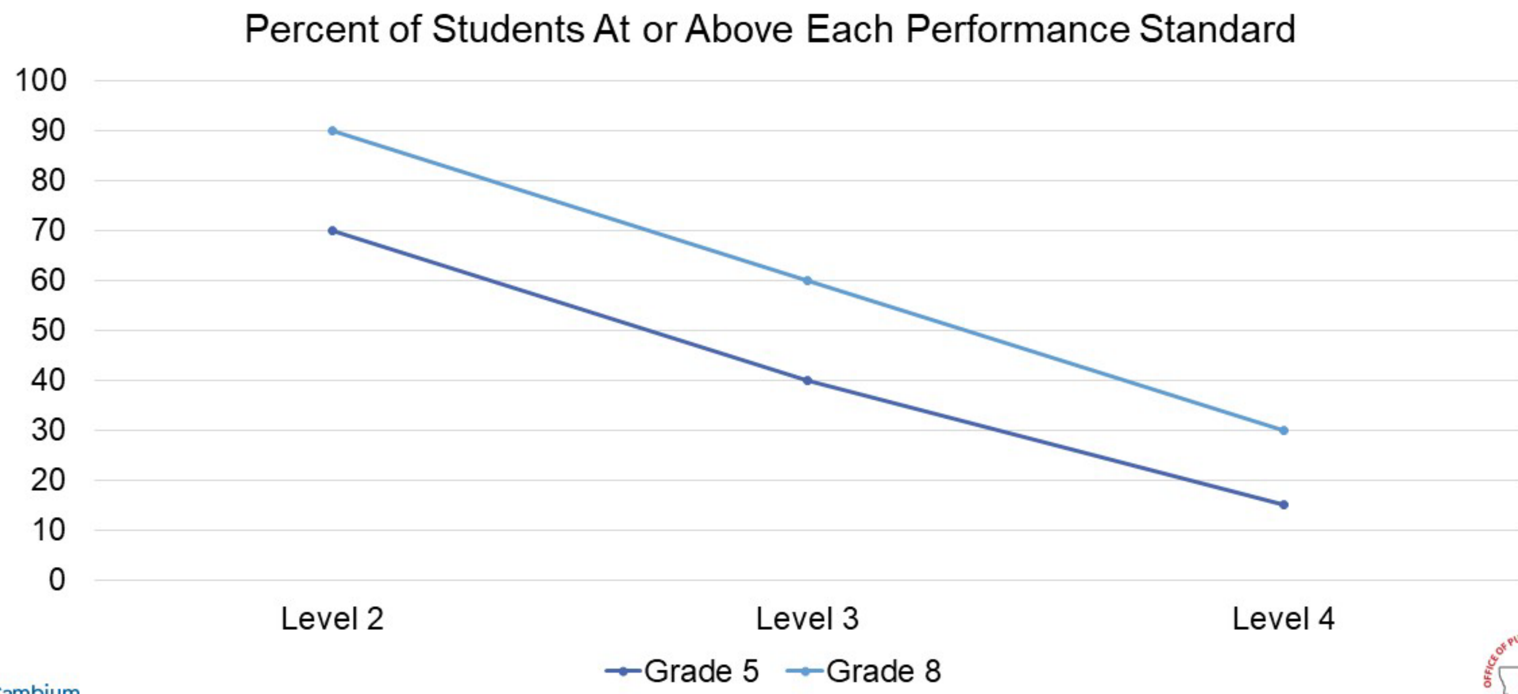
32

- Are performance standards nationally competitive and represent on track for college readiness?
  - NAEP Science
- Performance levels for benchmark assessments will provide context about the general neighborhood in which performance standards likely reside

2015 Montana NAEP Science Results			
Grade	At or Above		
	Basic	Proficient	Advanced
5	82	43	1
8	78	41	2

Articulation

33



Moderation

34

- After the standards have been recommended by the panelists, the Table Leaders meet to review the outcomes
  - All members are invited to observe this meeting but only the Table Leaders participate
- If there are anomalies across grades or subjects the Table Leaders are permitted to adjust the performance standards (assuming there is a good content reason for doing so)

Break Into Groups

35

Panel	Facilitators
Grade 5 Science	Hibbah Haddam Olivia Francois
Grade 8 Science	Matthew Davis Jared Taylor

Exhibit F-2. Breakout Room Slides

Standard Setting: Science

August 2 – 3, 2022
Montana Science Assessments (MSA)

2

Breakout Room Training

Cambium Assessment, Inc.

3

Welcome and Introductions

Workshop Leaders

4

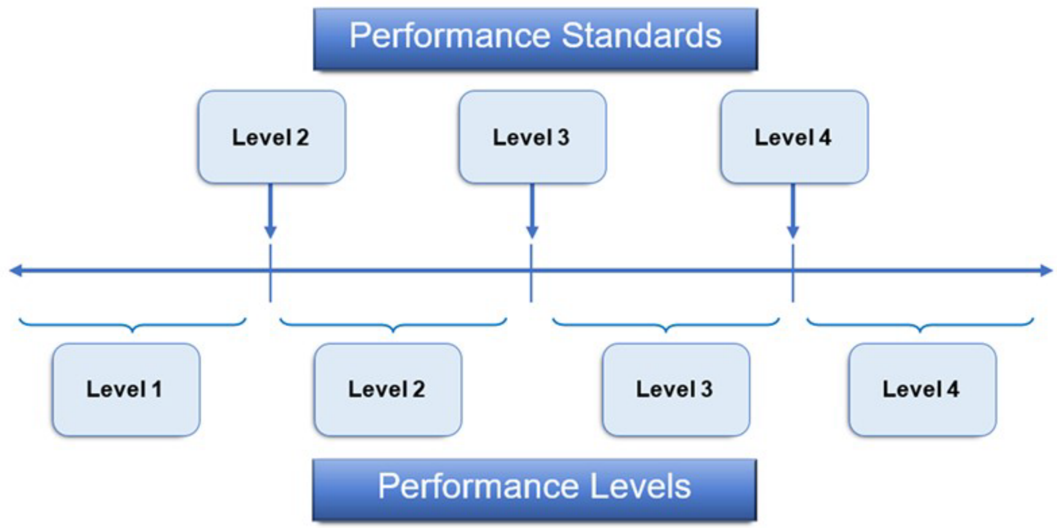
Cambium Assessment, Inc.

- Room Facilitators
  - Grade 5: Hibbah Haddam and Olivia Francois

Purpose of the Standard-Setting Workshop

5

- Recommend to the Montana Board of Public Education (BPE) three performance standards to differentiate the four performance levels on the MSA in grades 5 and 8



Cambium Assessment

Main Workshop Activities

6

- Large-Group Orientation
- Panel Training
  - Take the Online Operational Assessment
  - Review Range PLDs
  - Discuss Just Barely PLDs
  - Review the Ordered Scoring Assertion Booklet
  - Training on Assertion-Mapping Procedure
- Recommend Performance Standards
  - Two rounds
  - Panelist feedback following Round 1
  - Vertical Articulation
- Workshop Evaluation

Cambium Assessment

Importance of Security

7

Please do not:

- Create any form of electronic copy of test content (screenshots, electronic notes, etc.)
- Create any hand-written notes of test content
- Discuss test content with anyone outside the meeting
- Use your computer during the course of the meeting for any purpose other than participating in the item review (e.g., email, web browsing, social media)
- Save notes about item or passage content to your computer

Cambium Assessment

Reason for New Science Standards

8

- The Montana Board of Public Education adopted the Montana Science Content Standards in September 2016
- In spring 2022, assessments aligned to the 2016 Montana Science Content Standards were administered to all students in Montana in grades 5 and 8

Description of the Science Test Design

9

- Grades 5 and 8 tests assess students' understanding of the Montana Science Content Standards
- The MSA at each grade includes 6 item clusters and 12 stand-alone items
  - **Item clusters** include a stimulus and a series of questions that generally take students about 6–12 minutes to complete
  - **Stand-alone items** are shorter and generally take 1–3 minutes to complete
- All items ask students to use science and engineering practices and apply their understanding of disciplinary core ideas and crosscutting concepts to make sense out of real-world phenomena

Review of 3D Science Standards

10

- Each 3D "standard" is a blend of one or two "big ideas" from a science discipline (DCI), one of several scientific activities that are common to the doing of all science (SEP), and one of a number of broad themes that are found across scientific disciplinary boundaries (CCC).

Review of Items – 3D Composition

11

MS-LS1-1 From Molecules to Organisms: Structures and Processes

Students who demonstrate understanding can:

MS-LS1-1. Conduct an investigation to provide evidence that living things are made of cells; either one cell or many different numbers and types of cells. [Clarification Statement: Emphasis is on developing evidence that living things are made of cells, distinguishing between living and non-living things, and understanding that living things may be made of one cell or many and varied cells.]

The performance expectation above was developed using the following elements from the NRC document *A Framework for K-12 Science Education*:

Science and Engineering Practices	Disciplinary Core Ideas	Crosscutting Concepts
Planning and Carrying Out Investigations Planning and carrying out investigations in 6-8 builds on K-5 experiences and progresses to include investigations that use multiple variables and provide evidence to support explanations or solutions. • Conduct an investigation to produce data to serve as the basis for evidence that meet the goals of an investigation.	**LS1.A: Structure and Function** • All living things are made up of cells, which is the smallest unit that can be said to be alive. An organism may consist of one single cell (unicellular) or many different numbers and types of cells (multicellular).	**Scale, Proportion, and Quantity** • Phenomena that can be observed at one scale may not be observable at another scale. - ***Connections to Engineering, Technology and Applications of Science*** **Interdependence of Science, Engineering, and Technology** • Engineering advances have led to important discoveries in virtually every field of science, and scientific discoveries have led to the development of entire industries and engineered systems.

Review of Items – 3D Composition

12

- Three-dimensional science standards

Scientific and Engineering Practices	Crosscutting Concepts	Disciplinary Core Ideas
▸ Asking questions or defining problems ▸ Developing and using models ▸ Planning and carrying out investigations ▸ Analyzing and interpreting data ▸ Using mathematics and computational thinking ▸ Constructing explanations and designing solutions ▸ Engaging in argument from evidence ▸ Obtaining, evaluating, and communicating information	▸ Patterns ▸ Cause and effect: mechanism and explanation ▸ Scale, proportion, and quantity ▸ Systems and system models ▸ Energy and matter: flows, cycles, and conservation ▸ Structure and function ▸ Stability and change	▸ Earth and Space Science ▸ Life Science ▸ Physical Science ▸ Engineering

Item Clusters and Stand-Alone Items

13

- Item clusters
  - Designed to engage the student in grade-appropriate, meaningful scientific activity aligned to a specific standard
  - Item clusters include a stimulus and a series of questions that generally take students about 6–12 minutes to complete
- Stand-alone items are shorter and generally take students 1–3 minutes to complete

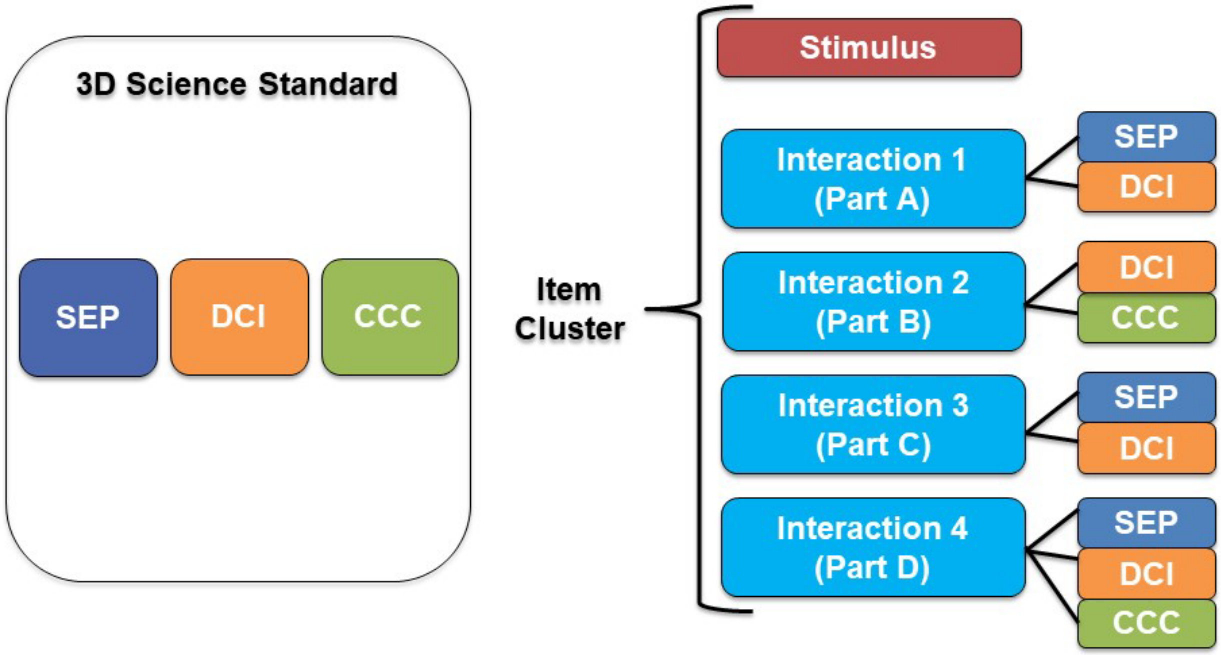
Structure of Item Clusters

14

- Each item cluster begins with a ***phenomenon***, which is the observation about the natural world which anchors the entire item cluster. The interactions within the item cluster all address the phenomenon.
- Each item cluster engages the student in a grade-appropriate, meaningful ***scientific activity*** aligned to a specific standard.
- A ***cluster task statement*** comes at the end of the stimulus and an overview of the point of the item cluster.
- Each measurable moment is captured with a ***scoring assertion***. These assertions clearly articulate what evidence the student has provided as a means to infer a specific skill or concept.

Review of Item Clusters – Composition

15



Review of Items – Composition Example

16

Sparks fly off the wheels of a train when the brakes are applied.

Click the small gray arrow to see a demonstration of this happening in Animation 1.

Animation 1. Braking Train

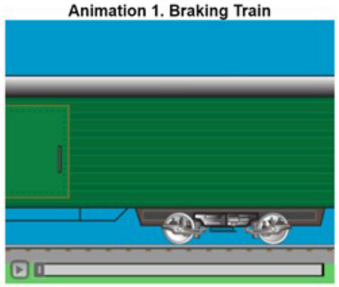


Table 1 explains some properties of the train and its surroundings as energy flows throughout the system.

Table 1. Properties of the Train System

Before Brakes Are Applied	After Brakes Applied
No sparks	Sparks fly off the wheels and brake pads
Brake pads make no sound	Brake pads make sound

1680

Part A

Click on each blank box to select the word or phrase that completes each sentence, constructing an argument about what happens when the train's brakes are applied.

Applying the brakes causes the [▾] to transfer kinetic energy to the [▾]. This causes the [▾] to slow down and have [▾] kinetic energy, which slows the train.

Part B

When the train applies its brakes, what happens to the energy of the surroundings?

(A) The surroundings gain energy.
(B) The surroundings lose energy.
(C) The surroundings do not gain or lose energy.
(D) There is not enough information to determine the energy of the surroundings.

Part C

Which **three** statements support your choice in part B?

- [ ] The train maintains its speed.
- [ ] Sound is produced.
- [ ] Sound is consumed.
- [ ] Light is produced.
- [ ] Light is consumed.
- [ ] Heat is produced.
- [ ] Heat is consumed.

Cambium Assessment

Scoring Assertions

17

- Within each item cluster, a series of explicit assertions can be made about the knowledge and skills that a student has demonstrated based on specific features of the student's responses
- Scoring assertions can be supported based on students' responses in one or more interactions within an item cluster.
- For example:
  - A student correctly graphs data points indicating that (s)he can construct a graph showing the relationship between two variables,

Review of Items – Scoring Assertions

18

Score Rationale	
The student selected "wheels" for the first blank and "brakes" or "rails" for the second blank showing an understanding of the interactions in the system and the effects of that energy flow.	✖
The student selected "wheels" for the third blank and "less" for the fourth blank showing an understanding of the interactions in the system and the effects of that energy flow.	✖
The student selected "The surroundings gain energy," showing an understanding of how the energy of the wheels change and is distributed throughout the system.	✖
The student selected "Sound is produced," providing evidence of how the energy of the surroundings has changed.	✖
The student selected "Light is produced," providing evidence of how the energy of the surroundings has changed.	✖
The student selected "Heat is produced," providing evidence of how the energy of the surroundings has changed.	✖
The student selected "The brakes make a screeching sound," which shows an understanding of how the energy changed throughout the system and that those changes serve as evidence that the the Kinetic Energy of the wheels transfers out of the wheels/system when the brakes are applied.	✖
The student selected "The sparks that fly off the wheels give off light," which shows an understanding of how the energy changed throughout the system and that those changes serve as evidence that the the Kinetic Energy of the wheels transfers out of the wheels/system when the brakes are applied.	✖
The student selected "The brakes give off energy as heat," which shows an understanding of how the energy changed throughout the system and that those changes serve as evidence that the the Kinetic Energy of the wheels transfers out of the wheels/system when the brakes are applied.	✖

Experience the Online Assessment

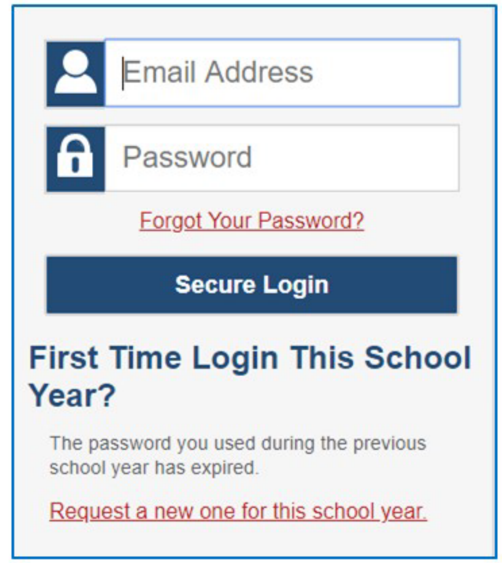
19

- Time to "Take the Test"
- Items administered in spring 2022
- Interface is similar to the online test environment that the students experienced
- This is an opportunity to interact with the items
- No need to "complete" the test, you will have more time later to become very familiar with the items
- You can score your responses
- You have ~90 minutes (stop at 11:15 am)

Accessing the Online Assessment

20

- Open the Chrome browser
- Sign in with your Username and Password



Cambium Assessment

21

Experience Online Operational Assessment

Step 2: Take the Operational Test

Cambium Assessment

OFFICE OF PUBLIC INSTRUCTION

Standard Setting

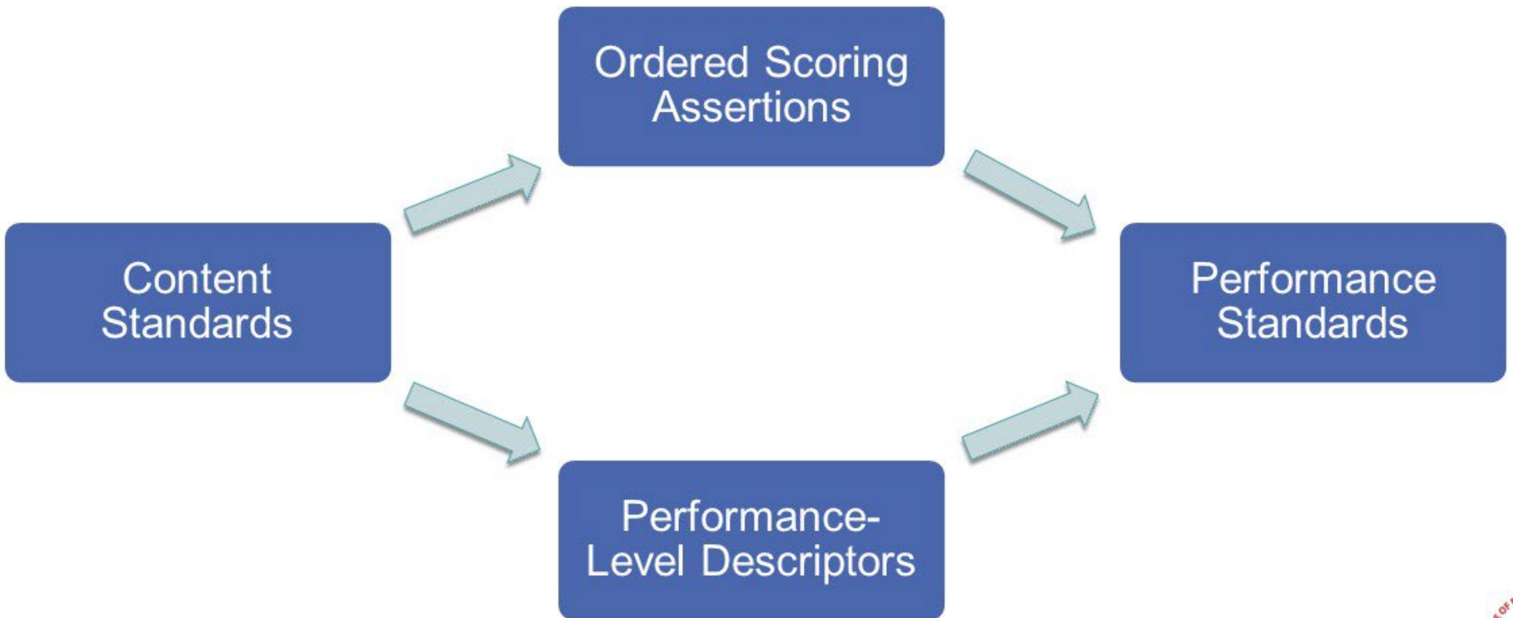
22

- Systematic process by which trained participants use their knowledge of academic content standards, test items, and student performance to recommend cut-scores associated with each performance level on the test

Cambium Assessment

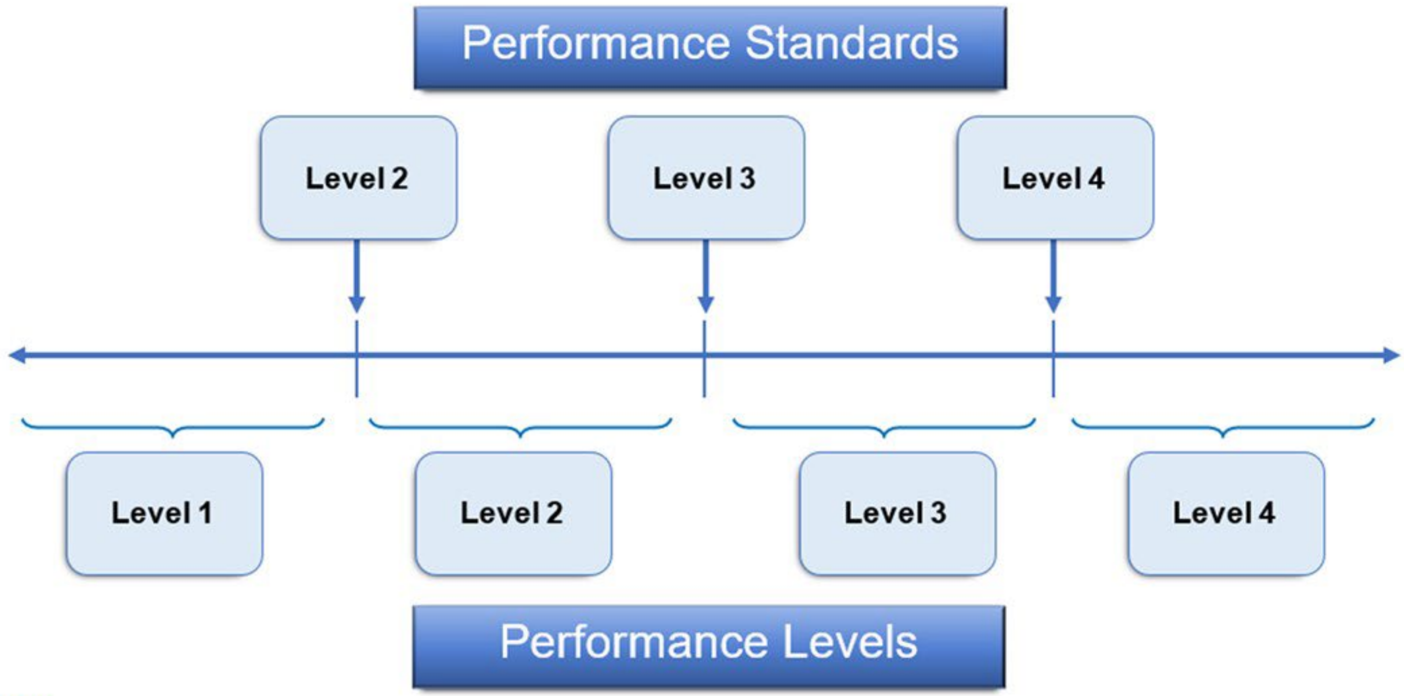
From Content Standards to Performance Standards

23



Performance Standards and Performance Levels

24



Assertion-Mapping Procedure (AMP)

25

- Test-centered procedure
- Employs an ordered item procedure adapted to accommodate new multiple interaction item types
- Map ordered scoring assertions to performance levels
- Is being employed to recommend performance standards in multiple states assessing three-dimensional science standards

Key Elements of the AMP

26

- Performance-level descriptors (PLDs)
  - Range PLDs
  - Threshold PLDs (just barely meets)
- Ordered scoring assertions
- Assertion map
- Assertion mapping in multiple rounds
  - Contextual information – benchmarking data and student impact data
  - Panelist feedback and group discussion
- Vertical articulation and moderation

Performance-Level Descriptors (PLDs)

27

- Describe what students within each performance level are expected to know and be able to do
- PLDs are the link between the content and performance standards

Grade 8 Range PLDs – Level 3

28

Physical Sciences

- **MS-PS1:** Develop and/or use a model to explain the conservation of mass when two substances react....
- **MS-PS2:** Ask questions, plan and conduct an investigation, and analyze and interpret data to make and support a claim regarding the relationship between mass, force, and motion, and the attractive and repulsive forces that act at a distance (electric, magnetic, and gravitational forces.
- **MS-PS3:** Develop and/or use a model or investigation to describe how kinetic and potential energy interact, transform, or transfer to another object…
- **MS-PS4:** Develop and/or use mathematical representations in a model to describe the patterns observed between wave characteristics and wave energy…

Cambium Assessment

Grade 8 Range PLDs Across Performance Levels

29

MS-PS3: Energy

- **Level 1: Identify components of a model** that investigates how kinetic and potential energy interact, transform, or transfer to another object; and **collect and record data for an investigation** that provides data regarding the temperature and total energy of a system and its dependency on a variety of factors, including the types and states of matter, as well as the amount of matter involved.
- **Level 2: Use a model to describe** how kinetic and potential energy interact, transform, or transfer to another object; and **collect and record data** regarding the temperature and total energy of a system and its dependence on a variety of factors, including the types and states of matter, as well as the amount of matter involved.
- **Level 3: Develop and/or use a model or investigation** to describe how kinetic and potential energy interact, transform, or transfer to another object; and **analyze data from an investigation to provide evidence** that the temperature and total energy of a system is dependent on a variety of factors, including the types and states of matter, as well as the amount of matter involved.
- **Level 4: Evaluate and/or revise a model to predict** changes to the interaction of kinetic and potential energy, including how energy is transformed, or transferred to another object; and **apply concepts of statistics and probability when providing evidence** that the temperature and total energy of a system is dependent on a variety of factors, including the types and states of matter, as well as the amount of matter involved.

Parse and Review the PLDs

30

- Take a few minutes to review the PLDs taking notice of the verbs and skills that differentiate the performance levels
  - Think about how the skills change from Below Basic to Advanced
  - Think about the skills and knowledge these students can demonstrate
  - Idea is to get a common mental representation of these students

  REMEMBER: Not every piece of content will be represented in the PLDs
- PLD Discussion

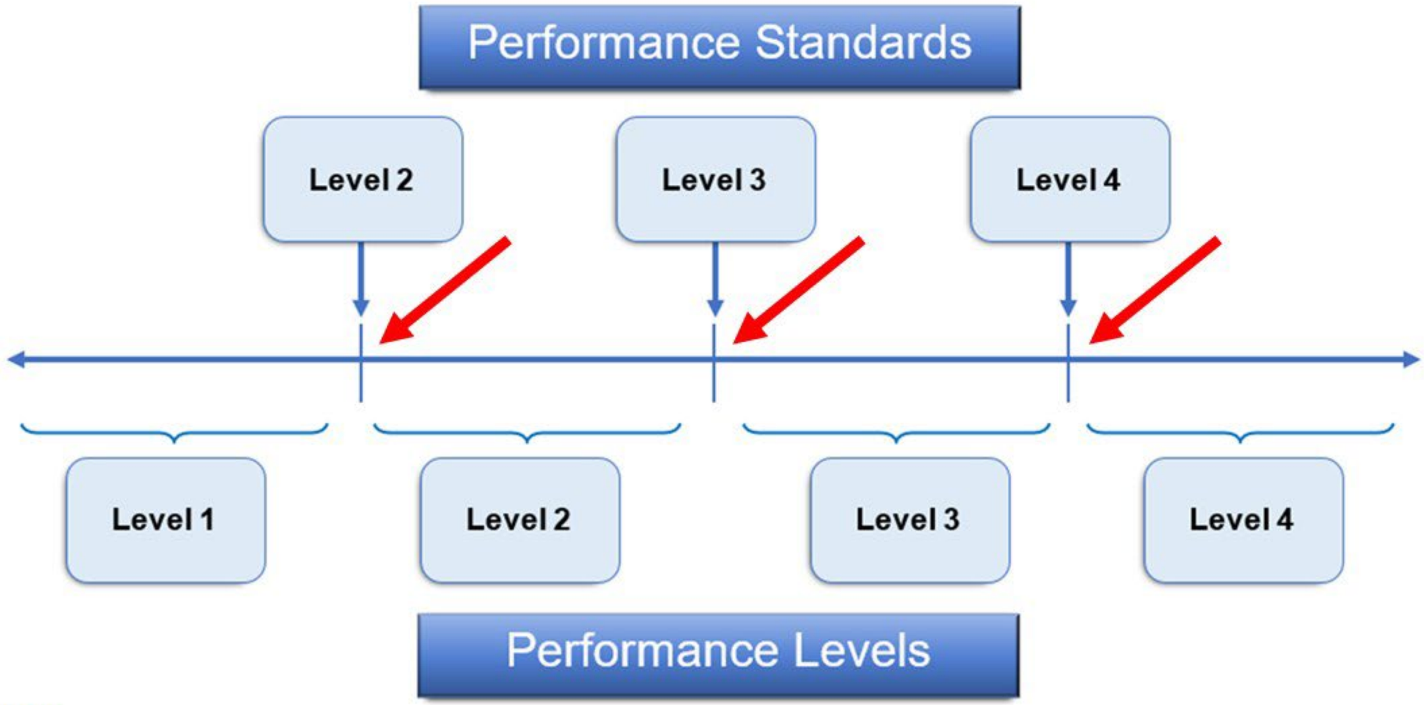
"Just Barely" Meets the Performance Standard

31

- When considering each performance level, we are especially interested in the transition areas between performance levels
- Pay attention to characteristics of students who ***just barely*** qualify for entry into the performance level from those just below
  - Not a typical example of students in the performance level
  - Although they are not good examples of the performance level, they do still meet the standard, or description in the PLD

Threshold "Just Barely" PLDs

32



Purpose of Just Barely Discussion

33

- Identify the types of skills these students can demonstrate
- Come to a common understanding of these skills and big ideas

Just Barely Discussion

34

- Think about what skills, concepts, or knowledge a just barely student would need to have to enter into each level
- As a group we will discuss the skills that a just barely student needs to have to gain entry into each of the four levels
- For each performance level think about:
  - What skills and knowledge must the student demonstrate to qualify for entrance into this performance level?
  - How does this differ from the upper range of the adjacent performance level?

Important Concepts

35

- "Just barely" meets the performance level
  - Differentiate students who just barely qualify for entry into a performance level from those just below
- Assertion mapping
  - Map each scoring assertion to the performance level that the assertion best supports
- Ordering of assertions
  - Assertions are ordered by difficulty within an item
  - Mapping of assertions to performance levels should reflect the ordering – no inversions within an item

36

Review of Ordered Scoring Assertion Booklet

Step 4: Review of Ordered Scoring Assertion Booklet

Ordered Scoring Assertions

37

- The ordered scoring assertion booklet (OSAB) constitutes a test administration:
  - A test form that meets test blueprint specifications
- It is important to evaluate scoring assertions as they relate to the item interactions
- Assertions within items are ordered by difficulty
  - Assertions within an item may not represent all PLDs

Cambium Assessment

What If an Assertion Seems Out of Order?

38

- Assertion ordering is based on student performance
- Assertions may seem out of order because they are ordered by difficulty, and not by content or cognitive process
- Identify why a scoring assertion is more difficult than the assertions before it, and easier than the assertions following it
  - Pay special attention to the interactions supporting the assertions
  - Assertions may be more or less difficult because of the underlying interactions

What If an Item Seems Wrong or Unfair?

39

- Do not let yourself get distracted – this is not an item review meeting
- If you believe something is wrong with an item interaction or scoring assertion, tell the Workshop Leader, then skip over the assertion as you review the rest of the assertions within the item

Cambium Assessment

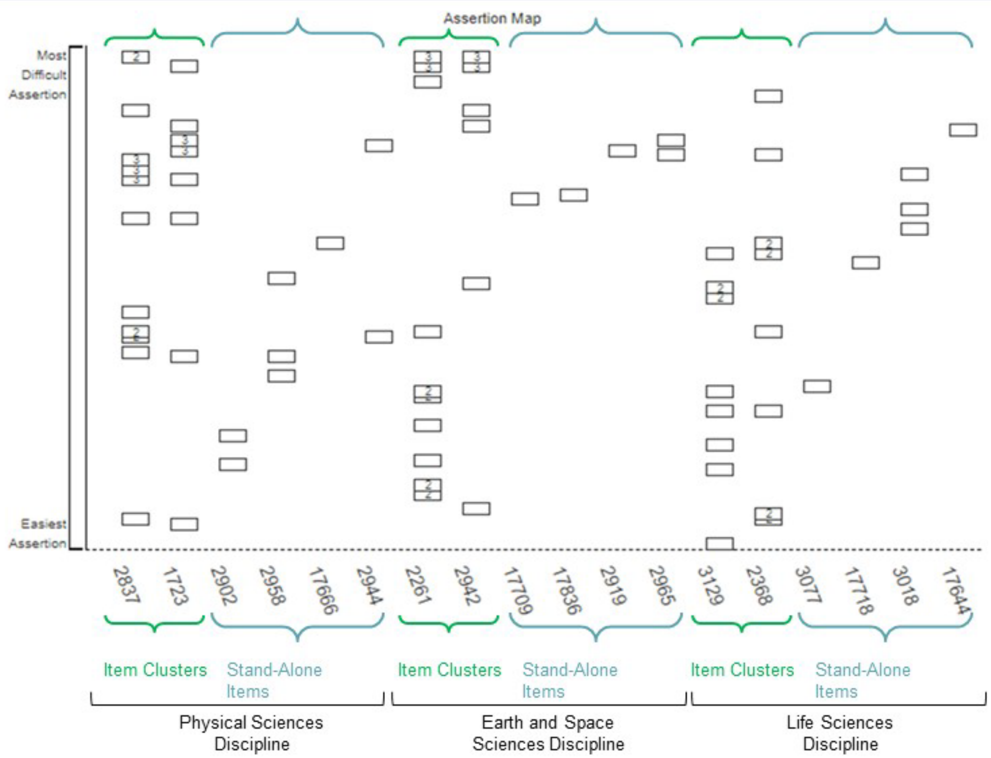
Studying the Items and Scoring Assertions

40

- Working individually, for each scoring assertion ask yourself:
  1. *How do the item interactions support the scoring assertion?*
  2. *Why is this assertion more difficult than the previous assertions (within the item)?*
  3. *How does the scoring assertion and the underlying interactions relate to the PLDs?*
- Working as a group
  - Discuss how item interactions support scoring assertions
  - Discuss ordering of scoring assertions
  - Discuss how scoring assertions are related to the PLDs

Assertion-Map

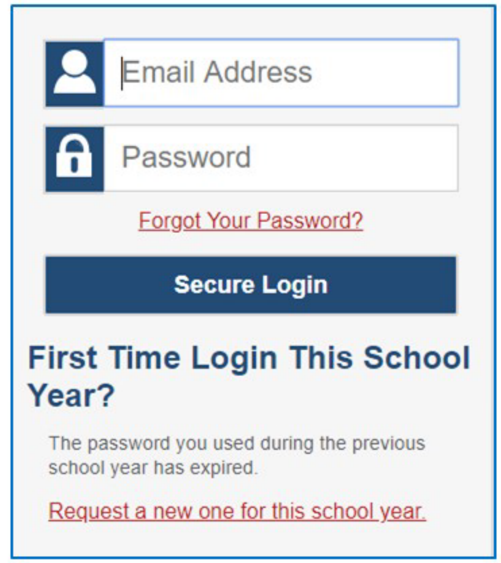
41



Accessing the OSAB

42

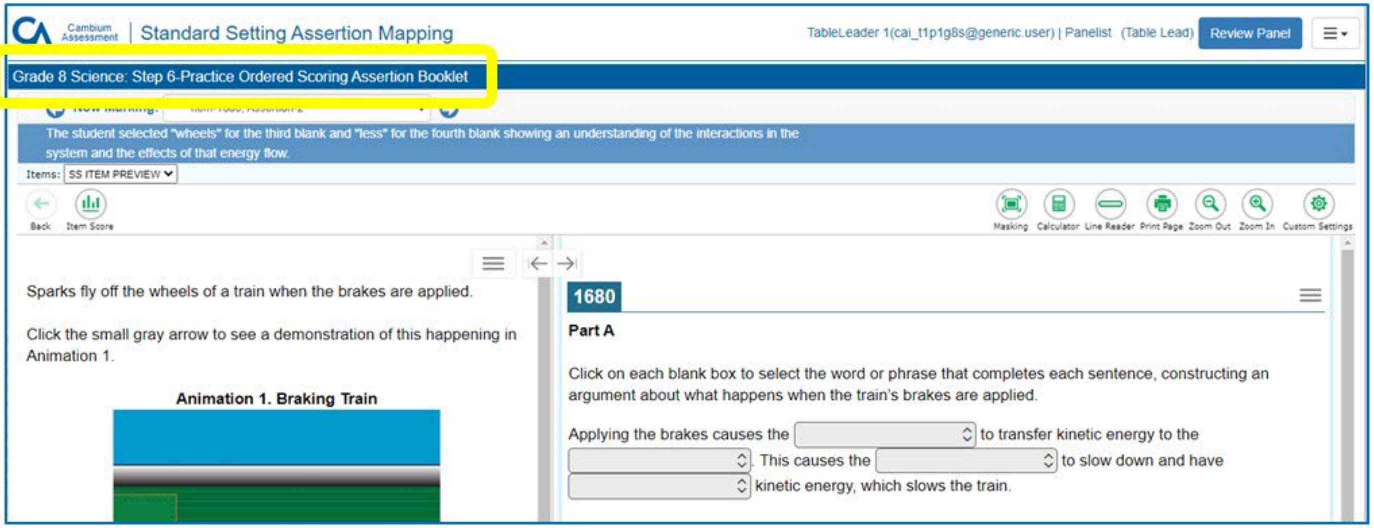
- Open the Chrome browser
- Sign in with your Username and Password



Navigating the OSAB

43

- Test and step we are working on shown at the top of the screen

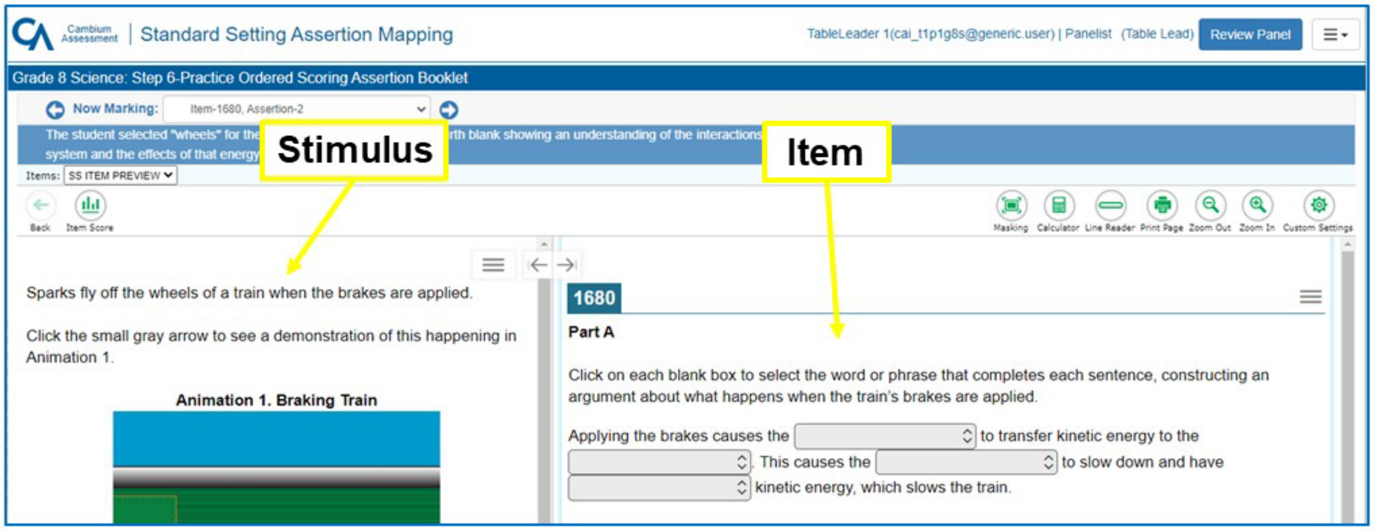


Cambium Assessment

Navigating the OSAB

44

- View the stimulus on the left side of the screen and the item on the right

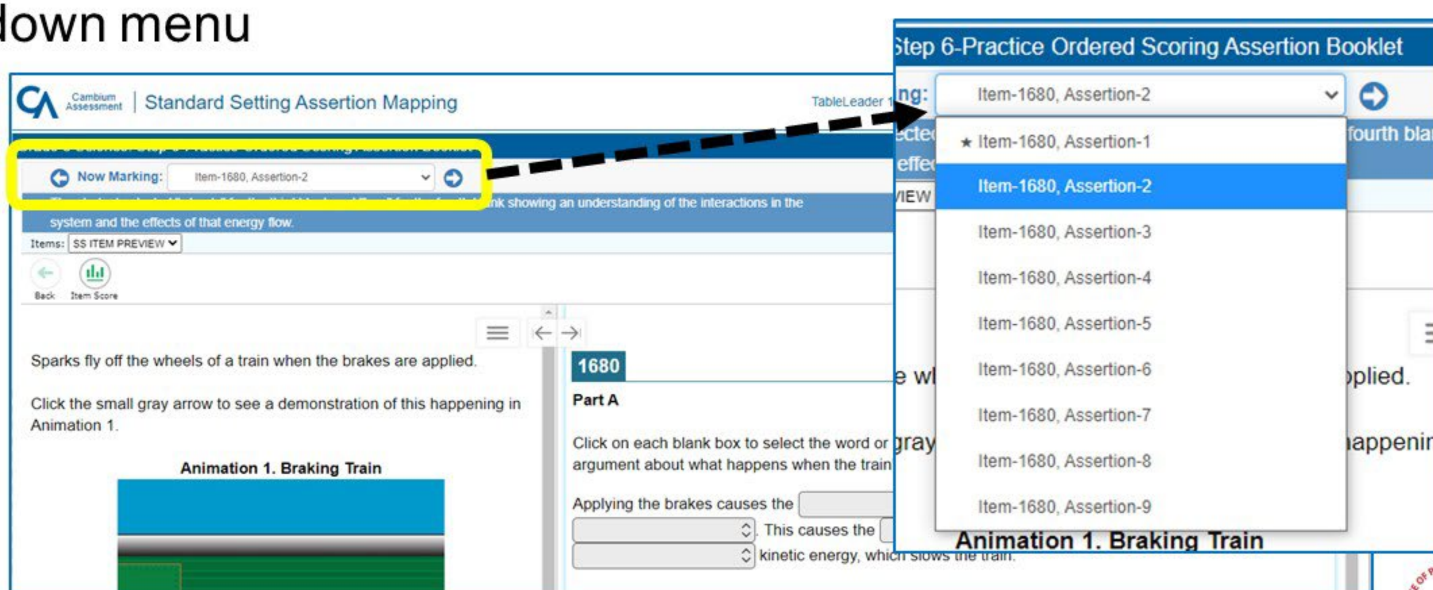


Cambium Assessment

Navigating the OSAB

45

- Move forward in the OSAB or select an assertion from the drop-down menu

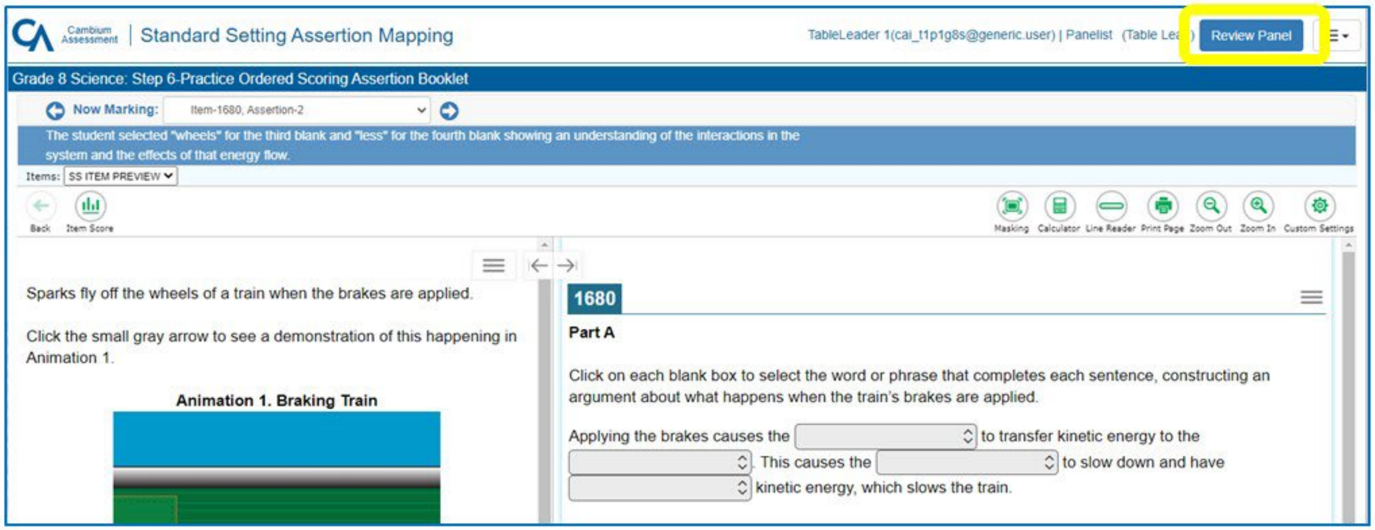


Cambium Assessment

Navigating the OSAB

46

- Access the Review Panel

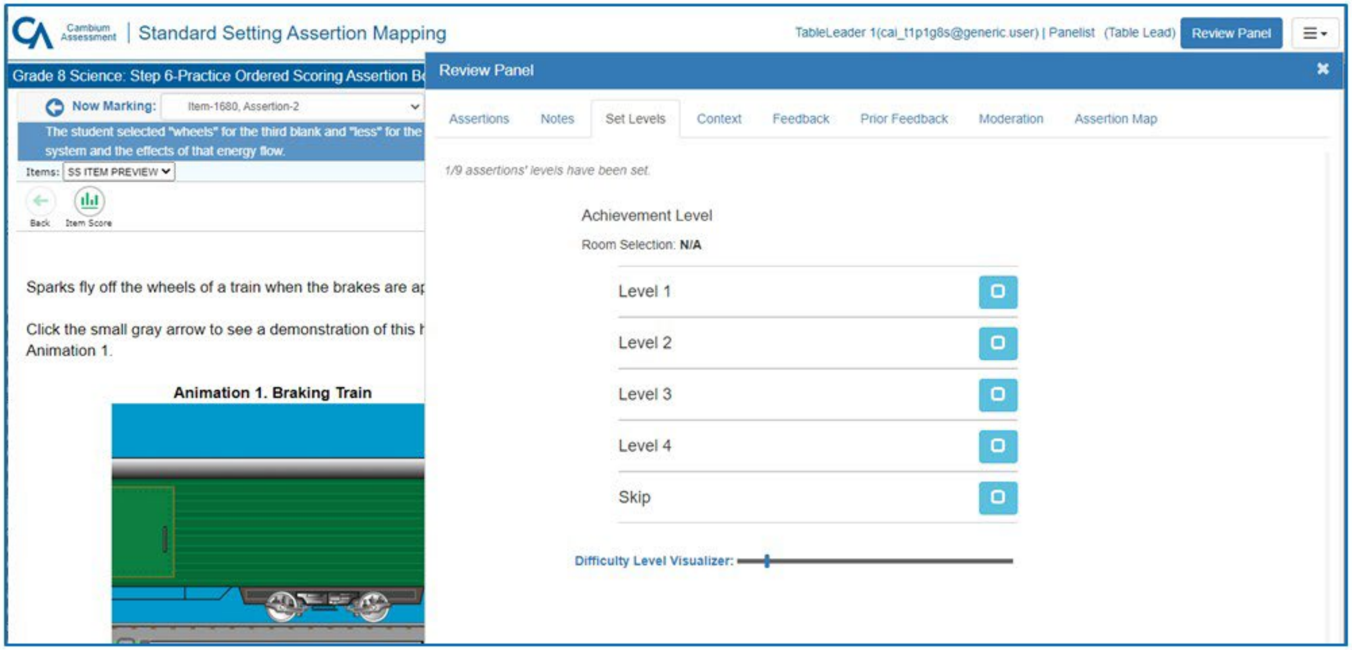


Navigating the OSAB – Review Panel

47

Navigating the OSAB – Review Panel

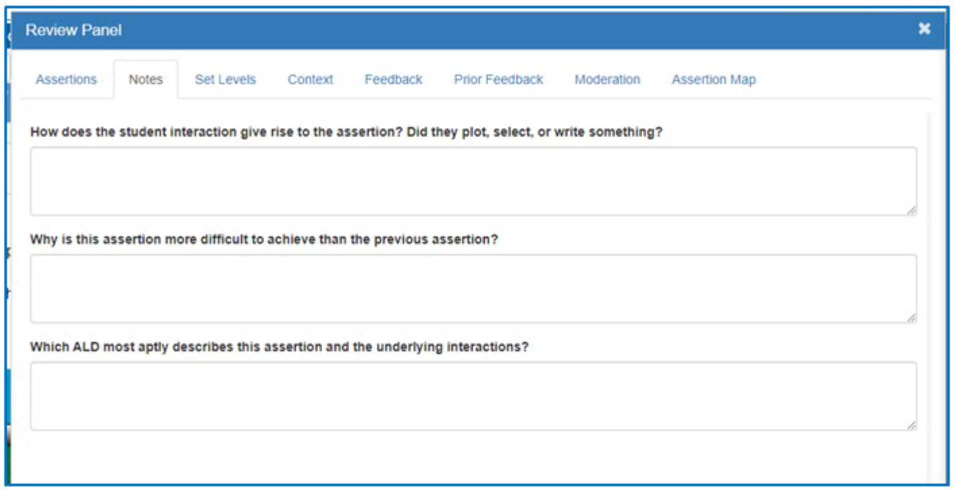
48



Navigating the OSAB – Review Panel

49

- “Context” tab – presents benchmarking and student impact data
- “Notes” tab – this is for your reference



Cambium Assessment

Studying the Items and Scoring Assertions

50

- We will work together on a set of items, asking and answering the following for each scoring assertion:
  1. *How do the item interactions support the scoring assertion?*
  2. *Why is this assertion more difficult than the previous assertions?*
  3. *How does the scoring assertion and the underlying interactions relate to the PLDs?*
- Then review the stand-alone items.

Benchmarking

51

- Are performance standards nationally competitive and represent on track for college readiness?
  - NAEP Science
- Performance levels for benchmark assessments will provide context about the general neighborhood in which performance standards likely reside

2015 Montana NAEP Science Results			
Grade	At or Above		
	Basic	Proficient	Advanced
5	82	43	1
8	78	41	2

Cambium Assessment

Assertion Map with Benchmark Information

52

Cambium Assessment

Review of the OSAB

53

- Let's review the items together

54

Standard Setting Day 2

Recommending Performance Standards for Grade 8 Science

Standard Setting Day 2 Agenda

55

- Continued review of OSAB
- Training on Assertion-Mapping Task
- Round 1 Assertion Mapping
- Review Feedback Data and Discuss Round 1 Results
- Round 2 Assertion Mapping
- Across Grade Moderation

Cambium Assessment

Continue review of OSAB

56

57

Training on Assertion-Mapping Task

Assertion-Mapping Task

58

- Map assertions to performance levels
  - Consider what differentiates students who just barely qualify for entry into the performance level from those not quite ready for entry into the performance level
  - Evidence that the student has demonstrated knowledge and skills necessary for entry into the performance level
- Map assertions in the online standard-setting tool

Benchmark and Impact Data

59

- Contextual information for each assertion
  - Benchmarking data: 2015 NAEP Science results
    - Shows the associated performance level for the NAEP Science Assessment if the Science standards were to exhibit the same rigor
  - Student impact data: percentage of students performing at or above the level associated with the assertion

Assertion Map with Benchmark Information

60

Mapping Ordered Assertions to Performance Levels

61

- You will map each scoring assertion to a performance level using the following tools:
  - PLDs
  - Difficulty Level Visualizer
  - Assertion Map
  - Contextual Information – benchmarking data and student impact data
  - Your professional judgement (and notes)
- Remember, scoring assertions are ordered from easiest to most difficult within each item
- **If you think that a subsequent assertion is at a lower level than a previous assertion, you might have been premature at mapping the level for the earlier assertion**
- You may "Skip" if an assertion seems to be out of place
  - Only use as a last resort

Cambium Assessment

62

Practice Assertion-Mapping Task and Standard-Setting Quiz

Step 6: Practice Ordered Scoring Assertion Booklet

Cambium Assessment

OFFICE OF PUBLIC INSTRUCTION

Practice Online Assertion-Mapping Task

63

- Purpose of this activity is to practice mapping assertions in the online environment. This is meant to help you become familiar with the tool and process.
  - Shortened version of the OSAB
  - One item cluster
- Log into the system and review the item cluster and ordered scoring assertions answering the three questions as you go
- Then, map each scoring assertion to a performance level and click "confirm"
- This is meant to help you become familiar with the tool and process

Practice quiz!

64

- Reminders for some key concepts

Round 1 Readiness Form

65

- Any questions?
- Is everyone ready for Round 1?
- If so, please fill out the readiness form

Round 1 Assertion Mapping

66

- You will map each assertion to a performance level
- Use the tools and documents along with your professional judgment
- Scoring assertions are ordered from easiest to most difficult within each item
- If you feel that a subsequent assertion is at a lower level than a previous assertion, then you might have been premature at mapping the level for the earlier assertion
- Should be a logical progress of performance levels (within an item)
  - No inversions
- You may “Skip” if, after consideration, the assertion seems to be out of place
  - Use as last resort
- When you have assigned all assertions click on the “Confirm” button
- This is an individual task
- Lunch is at 12:30 pm

Cambium Assessment

67

Round 1 Assertion Mapping

Step 8: Round 1 Assertion Mapping

Group Feedback and Discussion

68

- Goals
  - Add important information to your thinking
  - Develop common understandings
  - Inform possible re-evaluation of assertion mappings
- Expectation is converging judgments
  - Consensus is not a requirement or goal

Feedback and Impact Data

69

- Percentage of students reaching or exceeding the standard based on assertion mappings
- Group discussion
  - Does the percentage of students reaching or exceeding the current recommended performance standard seem reasonable?
  - What are the implications for the performance standards?
  - All performance standard recommendations should be based on content rationales drawn from the Montana Science Content Standards

Variance Monitor

70

- Consensus is NOT required, convergence is a goal
- Let's see where we have the most variance
- Discuss within each table for 15 minutes
- Then, we will come together for group conversation for 15 minutes

Variance Monitor

71

Lowest Variance → Highest Variance

Item Assertion Id	Difficulty Level	Mean	Room - 1 | Table - 1				Room - 1 | Table - 2				
			VI_T1P3_G5S	VI_T1P4_G5S	VI_T1P5_G5S	VI_T1P7_G5S	VI_T2P2_G5S	VI_T2P3_G5S	VI_T2P4_G5S	VI_T2P6_G5S	VI_T2P7_G5S
18027-Assertion-7	505	1.89	2	2	3	1	3	3	1	1	1
18027-Assertion-8	517	2.11	2	2	3	1	3	3	2	2	1
18027-Assertion-1	525	2.33	2	2	3	2	3	3	2	2	2
18027-Assertion-5	526	2.38	–	2	3	2	3	3	2	2	2
18027-Assertion-3	535	2.67	3	2	3	2	3	3	2	3	3
18027-Assertion-9	545	3.11	3	3	3	3	3	3	3	4	3
18027-Assertion-2	550	3.11	3	3	3	3	3	3	3	4	3
18027-Assertion-6	552	3.22	3	4	3	3	3	3	3	4	3
2931-Assertion-1	518	1.89	3	1	1	2	2	3	1	2	2

72

Round 2 Assertion Mapping

Step 12: Round 2 Assertion Mapping

Round 2 Readiness Form

73

- Any questions?
- Is everyone ready for Round 2?
- If so, please fill out the readiness form

Round 2 Assertion Mapping

74

- You will use the next hour to map each scoring assertion to a performance level
- Use the tools and documents along with your professional judgment, **contextual information – benchmark ad impact data**, and **feedback data**
- Scoring assertions are ordered from easiest to most difficult within each item
- If you feel that a subsequent assertion is at a lower level than a previous assertion, then you might have been premature at setting the level for the earlier assertion
    - No inversions
- You may “Skip” if, after consideration, the assertion seems to be out of place
    - Use as a last resort
- When you have assigned all assertions click on the “Confirm” button
- This is an individual task
- You have until 3:30 pm
- **Complete evaluations**

Cambium Assessment

75

Round 2 Results

Step 14: Results of Round 2

76

Moderation

Step 16: Moderation

Cambium Assessment

Creating a System of Performance Standards

77

- Performance standards for a statewide system must be coherent across grades and subjects
  - Articulation
  - Benchmarking
  - Moderation

Benchmarking

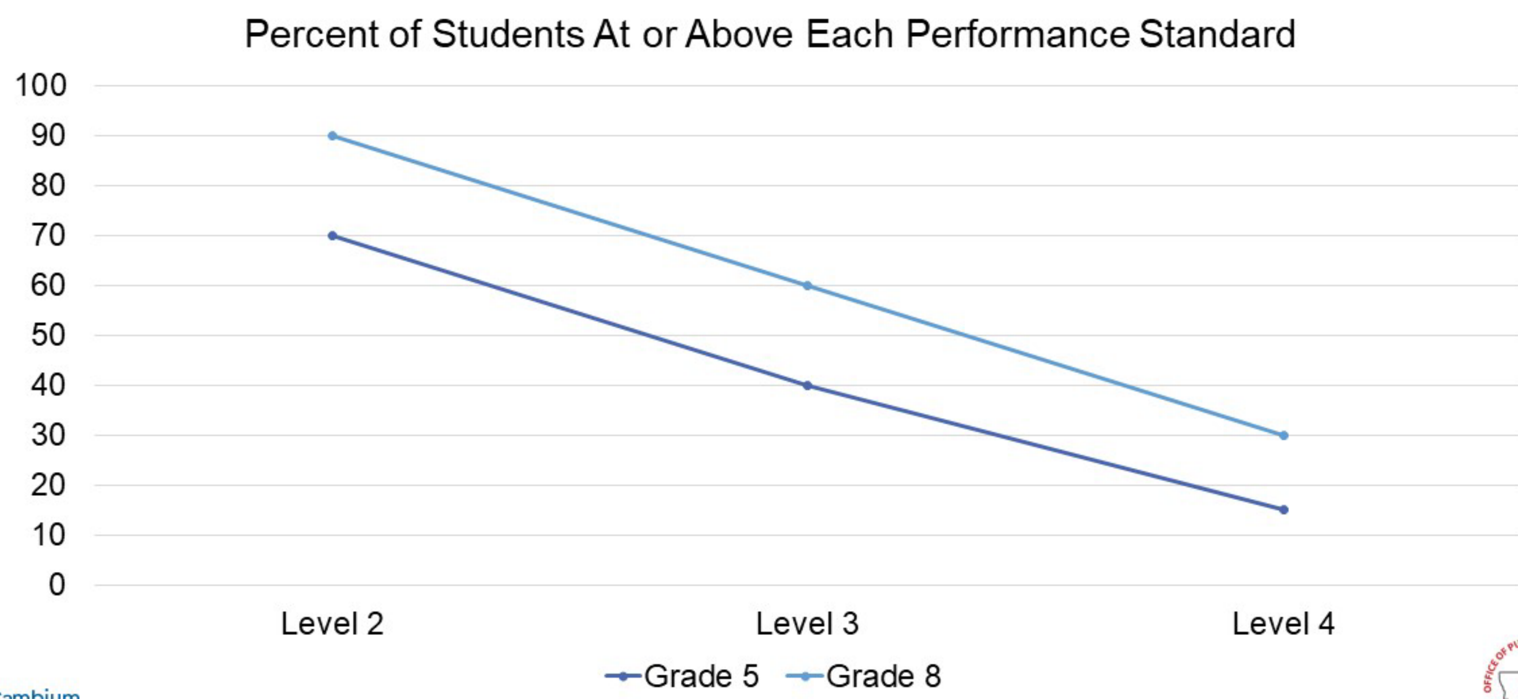
78

- Are performance standards nationally competitive and represent on track for college readiness?
  - NAEP Science
- Performance levels for benchmark assessments will provide context about the general neighborhood in which performance standards likely reside

2015 Montana NAEP Science Results			
Grade	At or Above		
	Basic	Proficient	Advanced
5	82	43	1
8	78	41	2

Articulation

79



Moderation

80

- After the standards have been recommended by the panelists, the Table Leaders meet to review the outcomes
  - All members are invited to observe this meeting but only the Table Leaders participate
- If there are anomalies across grades or subjects the Table Leaders are permitted to adjust the performance standards (assuming there is a good content reason for doing so)

Appendix 3-G

Standard-Setting Practice Quiz

Standard-Setting Practice Quiz

Exhibit G-1. Standard-Setting Practice Quiz

2022 Montana Science Assessment Standard Setting - Assertion Mapping Practice Quiz

* Required

1. Name: *

   ____

2. Panelist ID (e.g., MT_T1P1_G5S): *

   ____

3. Assigned Committee: *

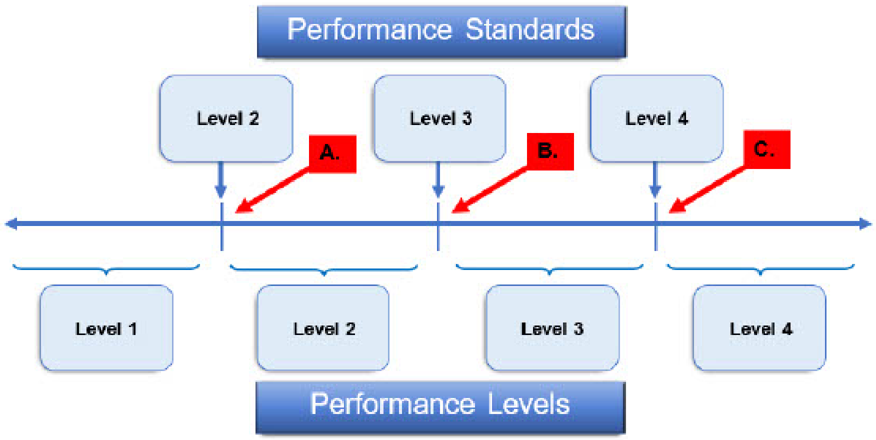
   Mark only one oval.

   ◯ Science Grade 5

   ◯ Science Grade 8

Performance Standards and Performance Levels

The graphic below illustrates the relationship between the performance standards that you will recommend and the performance levels that they demarcate:



4. Which red box on the performance continuum graphic above illustrates students who are just barely described by the Level 3 PLD? *

Mark only one oval.

- Box A
- Box B
- Box C

5. Which red box on the performance continuum graphic above illustrates students who are just barely described by the Level 2 PLD? *

Mark only one oval.

- Box A
- Box B
- Box C

6. Which red box on the performance continuum graphic above illustrates students who are just barely described by the Level 4 PLD? *

   Mark only one oval.

   - Box A
   - Box B
   - Box C

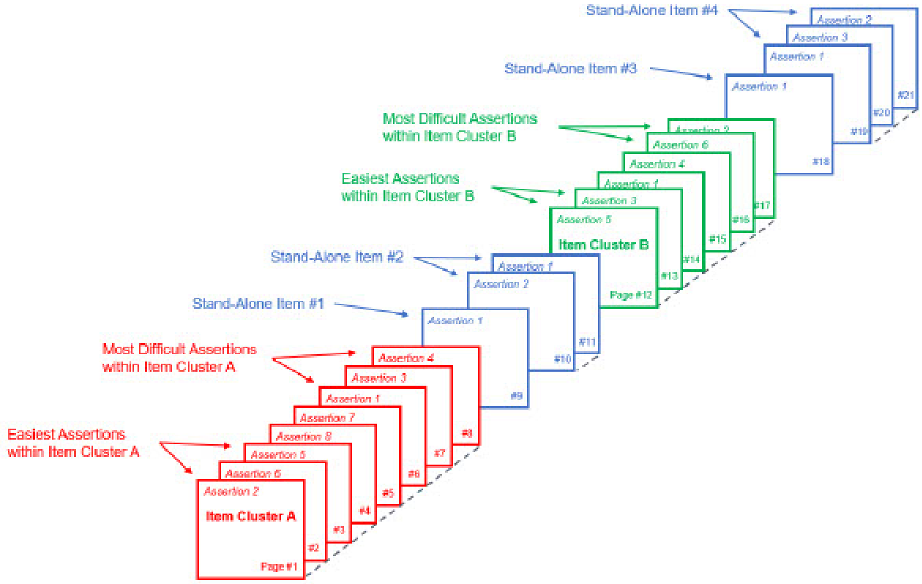
7. Which performance standard differentiates between the Level 2 performance level and the Level 3 performance level? *

   Mark only one oval.

   - Level 2
   - Level 3
   - Level 4

Ordered Scoring Assertion Booklet (OSAB)

Here is a hypothetical Ordered Scoring Assertion Booklet (OSAB) that consists of pages 1 through 21:



8. Within each item cluster or stand-alone item within the OSAB, scoring assertions are ordered by difficulty. In the hypothetical OSAB presented above, is the assertion on page 7 of the OSAB easier, more difficult, or about the same as the assertion on page 3? *

Mark only one oval.

- The assertion on page 7 is easier than the assertion on page 3
- The assertion on page 7 is more difficult than the assertion on page 3
- The assertion on page 7 is about the same as the assertion on page 3
- The difficulty of the assertions on pages 7 and 3 cannot be compared in this graphic because they are not within the same item

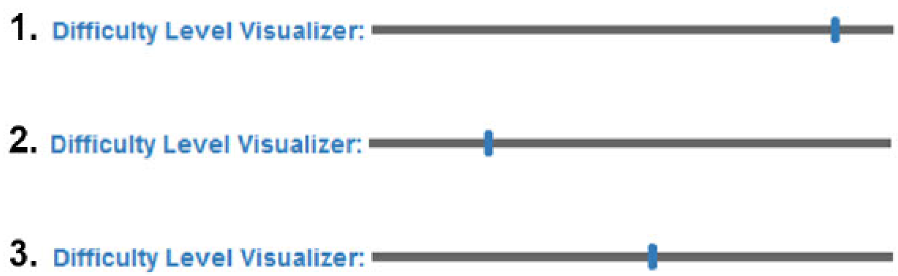
Standard-Setting Assertion Mapping Tool

9. Do you have to assign each scoring assertion to a performance level (or use the skip button)? *

   Mark only one oval.

   - ◯ Yes
   - ◯ No

Below are three different scoring assertions' Difficulty Level Visualizers.



10. Which Difficulty Level Visualizer in the image above represents the most difficult scoring assertion? *

    Mark only one oval.

    - ◯ Difficulty Level Visualizer 1
    - ◯ Difficulty Level Visualizer 2
    - ◯ Difficulty Level Visualizer 3

11. Which Difficulty Level Visualizer in the image above represents the least difficult scoring assertion? *

Mark only one oval.

- Difficulty Visualizer 1
- Difficulty Visualizer 2
- Difficulty Visualizer 3

Appendix 3-H

Standard-Setting Readiness Forms

Standard-Setting Readiness Forms

Exhibit H-1. Standard-Setting Round 1 Readiness Form

2022 Montana Science Assessment Standard Setting Educator Panel - Readiness Form

Preparation for Round 1 Assertion Mapping

* Required

1. Name: *

   ____

2. Panelist ID (e.g., MT_T1P1_G5S): *

   ____

3. Assigned Committee: *

   Mark only one oval.

   ◯ Science Grade 5
   ◯ Science Grade 8

Preparation for Round 1 Assertion Mapping

4. The workshop training has prepared me to review the Performance-Level Descriptors (PLDs) and fully explained the concept of threshold PLDs. *

   Mark only one oval.

   ◯ Yes
   ◯ No

5. The workshop training has prepared me to review the Ordered Scoring Assertion Booklet (OSAB). *

Mark only one oval.

- Yes
- No

6. The workshop training has clearly explained how to use the assertion map when reviewing the OSAB. *

Mark only one oval.

- Yes
- No

7. The workshop training has clearly explained the task of mapping assertions in the OSAB to the performance levels in the standard-setting tool. *

Mark only one oval.

- Yes
- No

8. The workshop training has fully explained how to use the contextual information (student impact data and benchmarking data) when mapping assertions to performance levels. *

Mark only one oval.

- Yes
- No

9. I have answered "Yes" to the above questions and I understand what I need to do to map assertions to performance levels. (Please initial below.) *

Mark only one oval.

- ⬭ Yes
- ⬭ No

10. Initial: *

____

11. If I answered "No" to any of the above questions, I received additional training. (Please initial below.) *

Mark only one oval.

- ⬭ Yes
- ⬭ No
- ⬭ Not applicable

12. Initial: *

____

13. Following the additional training, I feel sufficiently trained on what I need to do to map assertions to performance levels. (Please initial below.) *

Mark only one oval.

- ⬭ Yes
- ⬭ No
- ⬭ Not applicable

14. **Initial:** *

This content is neither created nor endorsed by Google.

Google Forms

Exhibit H-2. Standard-Setting Round 2 Readiness Form

2022 Montana Science Assessment Standard Setting Educator Panel - Readiness Form

Preparation for Round 2 Assertion Mapping

* Required

1. Name: *

   ____

2. Panelist ID (e.g., MT_T1P1_G5S): *

   ____

3. Assigned Committee: *

   Mark only one oval.

   ◯ Science Grade 5
   ◯ Science Grade 8

Preparation for Round 2 Assertion Mapping

4. The workshop training has clearly explained how to use the assertion map when reviewing the Ordered Scoring Assertion Booklet (OSAB). *

   Mark only one oval.

   ◯ Yes
   ◯ No

5. The workshop training has clearly explained the task of mapping assertions in the OSAB to the performance levels in the standard-setting tool. *

   Mark only one oval.

   - Yes
   - No

6. The workshop training has fully explained how to use the contextual information (student impact data and benchmarking data) when mapping assertions to performance levels. *

   Mark only one oval.

   - Yes
   - No

7. The training fully explained the panel feedback data and impact data that was presented. *

   Mark only one oval.

   - Yes
   - No

8. I understand my task for Round 2. *

   Mark only one oval.

   - Yes
   - No

9. I have answered "Yes" to the above questions and I understand what I need to do to map assertions to performance levels. (Please initial below.) *

   Mark only one oval.

   - Yes
   - No

10. Initial: *

    ____

11. If I answered "No" to any of the above questions, I received additional training. (Please initial below.) *

    Mark only one oval.

    - Yes
    - No
    - Not applicable

12. Initial: *

    ____

13. Following the additional training, I feel sufficiently trained on what I need to do to map assertions to performance levels. (Please initial below.) *

    Mark only one oval.

    - Yes
    - No
    - Not applicable

14. **Initial:** *

____

Appendix 3-I

Round 1 and Round 2 Standard-Setting Assertion Maps

Round 1 Standard-Setting Assertion Maps

Exhibit I-1. Round 1 Standard-Setting Assertion Map, Science Grade 5

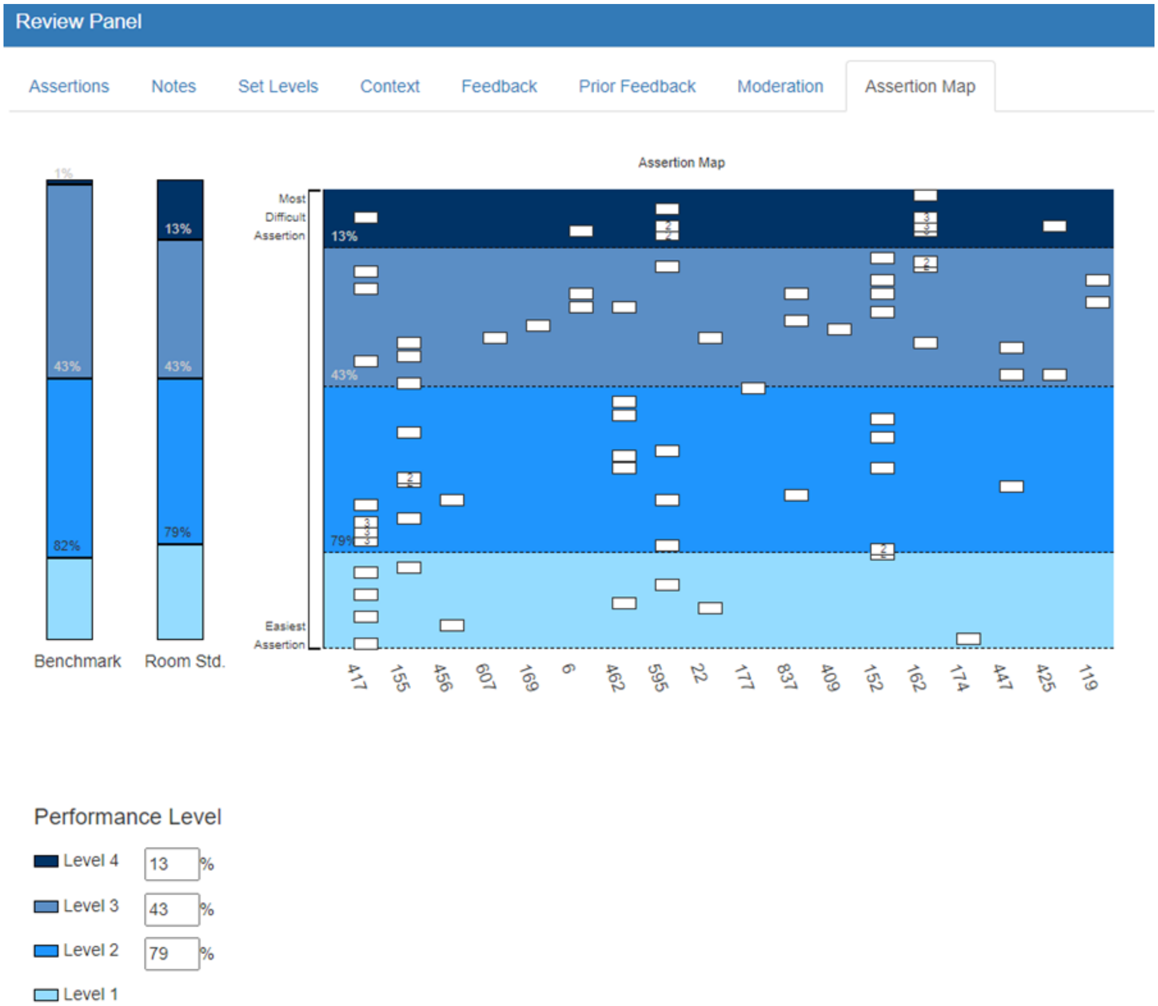
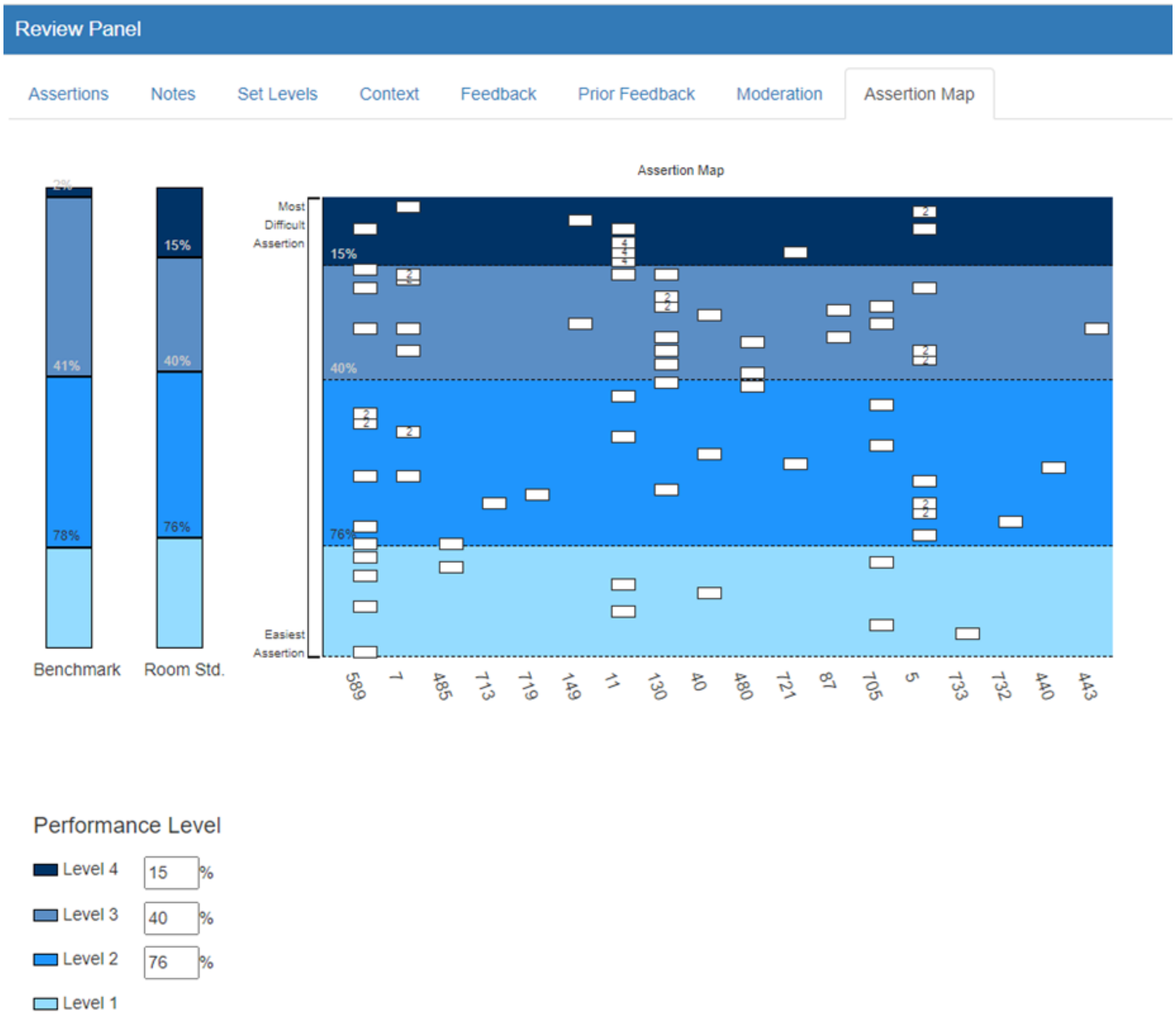


Exhibit I-2. Round 1 Standard-Setting Assertion Map, Science Grade 8



Round 2 Standard-Setting Assertion Maps

Exhibit I-3. Round 2 Standard-Setting Assertion Map, Science Grade 5

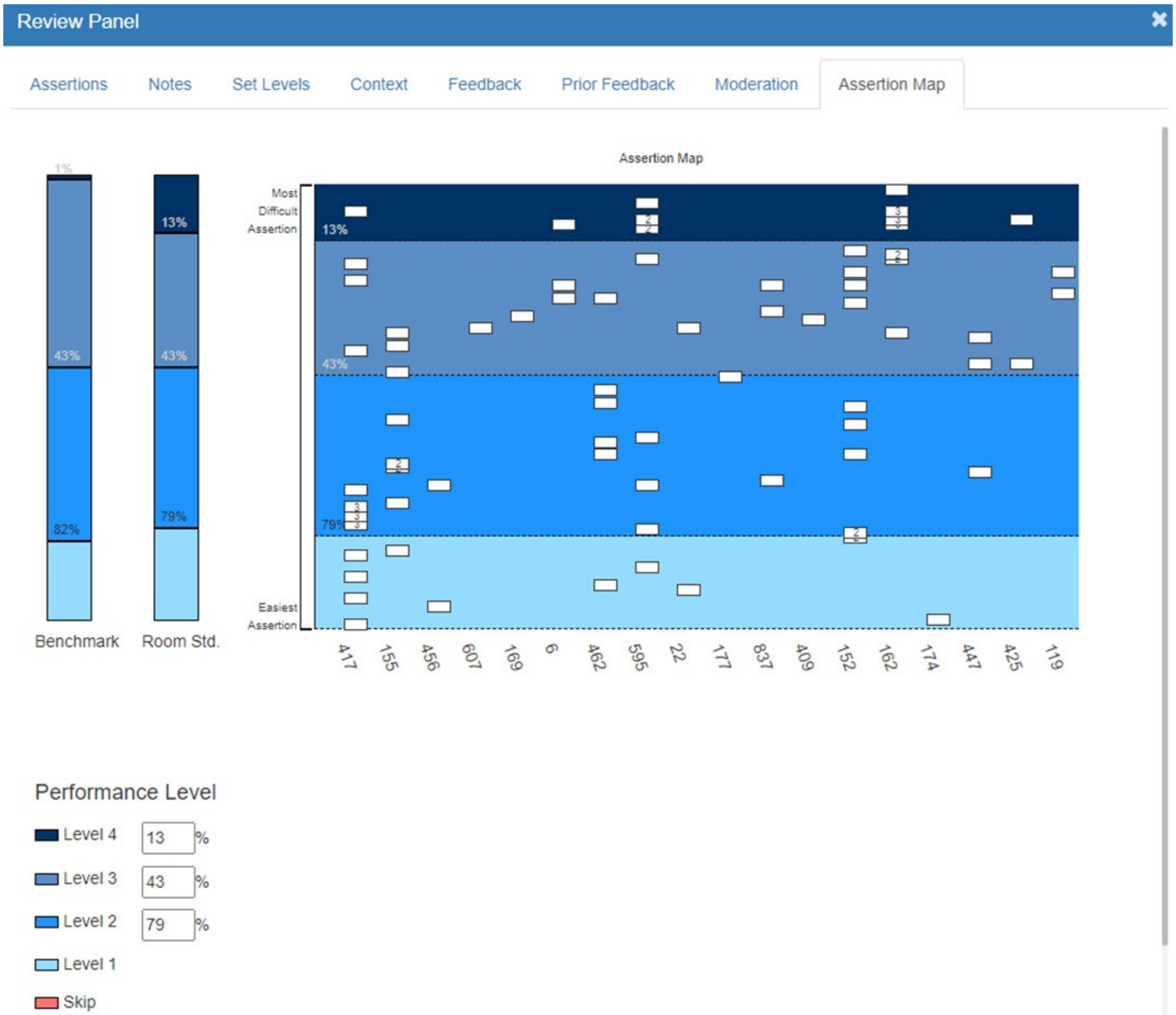
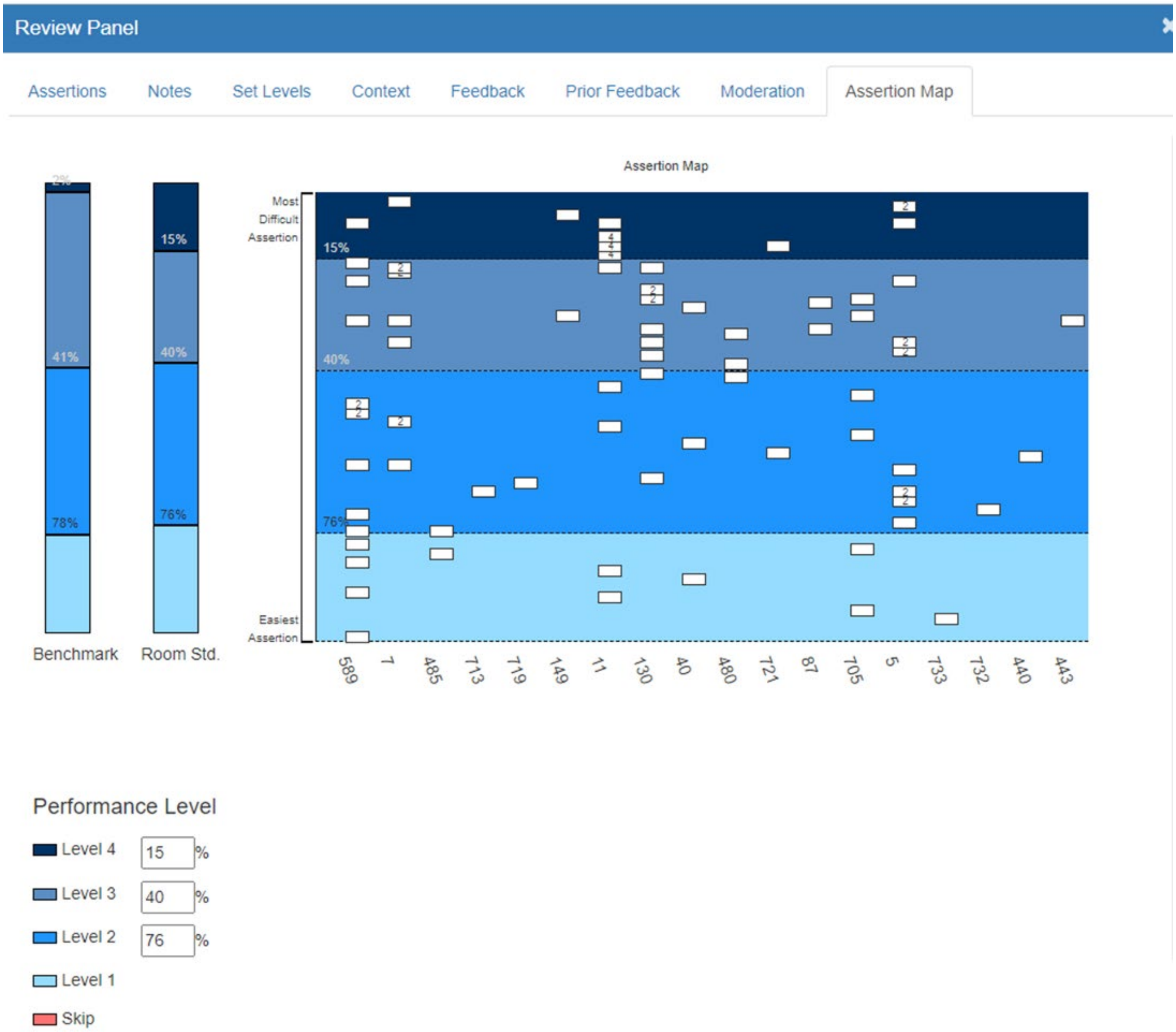


Exhibit I-4. Round 2 Standard-Setting Assertion Map, Science Grade 8



Montana Science Assessment

2022–2023

Volume 4:
Evidence of Reliability and Validity

OFFICE OF PUBLIC INSTRUCTION
ELSIE ARNTZEN, STATE SUPERINTENDENT

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

1. INTRODUCTION AND OVERVIEW OF RELIABILITY AND VALIDITY EVIDENCE

The state of Montana first implemented the Montana Science Assessment (MSA) for operational use in the 2021–2022 school year. The MSA is administered online to students in grades 5 and 8 using an adaptive test design. Accommodated versions are available at each grade level, including braille and Data Entry Interface (DEI) forms. Spanish-language versions of the tests are also available. Table 1 shows the complete list of tests that were available during the operational test administration in spring 2023.

Table 1. Spring 2023 Assessment Modes

Language/Format	Assessment Mode	Grade
English	Online	5 and 8
Spanish	Online	5 and 8
English Braille	Online	5 and 8
English Braille with Data Entry Interface (DEI)	Paper*	5 and 8

Note. *This refers to an online submission of the paper-pencil test. All student responses were entered into the Data Entry Interface (DEI) by test administrators.

Given the intended uses of these tests, both reliability evidence and validity evidence are necessary to support appropriate inferences of student academic performance from the MSA scores. The analyses to support reliability and validity evidence that are discussed in this volume were conducted on the basis of test scores reported for the online English-language version and the accommodated versions of the MSA.

The purpose of this report is to provide empirical evidence that can subsequently be used to support a validity argument for the uses of and inferences from the MSA. This volume addresses the following five topics:

- ***Reliability.*** The reliability estimates are presented by grade and demographic subgroup. This section also includes conditional standard errors of measurement (CSEM) and classification accuracy and consistency results by grade.
- ***Content Validity.*** This section presents evidence showing that all students' tests were constructed to measure the Montana Science Content Standards, with a sufficient number of items targeting each area of the test blueprint.
- ***Internal Structure Validity.*** Evidence is provided regarding the internal relationships among the subscale scores to support their use and to justify the item response theory (IRT) measurement model. This type of evidence includes observed and disattenuated Pearson correlations among discipline scores per grade. As explained in detail in Volume 1, Annual Technical Report, the IRT model is a multidimensional model, with an overall dimension representing proficiency in science and nuisance dimensions that consider within-item local dependencies among scoring assertions. In this volume, evidence is provided with

respect to the presence of item cluster effects. Additionally, confirmatory factor analysis (CFA) was used to evaluate the fit of the IRT model and to compare it with alternative models, including models with a simpler internal structure (e.g., unidimensional models) and models with a more elaborate internal structure.

- ***Relationship of Test Scores to External Variables.*** In this section evidence of convergent and discriminant validity is presented using observed and disattenuated subscore correlations both within and across subjects.
- ***Test Fairness.*** Fairness is an explicit concern during item development. Items are developed following the principles of universal design (UD), which. Universal design removes barriers in order to provide access for the widest possible range of students. Test fairness is further monitored statistically using differential item functioning (DIF) analysis in tandem with content reviews by specialists.

1.1 Reliability

The term *reliability* refers to consistency in test scores. Reliability can be defined as the degree to which individuals' deviation scores remain relatively consistent over repeated administrations of the same test or alternate test forms (Crocker & Algina, 1986). For example, if a person repeatedly takes the same or parallel tests, they should receive consistent results. The reliability coefficient refers to the ratio of true score variance to observed score variance:

$$\rho_{XX\prime} = \frac{\sigma_T^2}{\sigma_X^2}.$$

Another way to view reliability is to consider its relationship with the standard errors of measurement (SEM)—the smaller the standard error, the higher the precision of the test scores. For example, classical test theory (CTT) assumes that an individual's observed score (X) can be expressed as a true score (T) plus some error (E), $X = T + E$. The variance of X can be shown to be the sum of two orthogonal variance components:

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

Returning to the definition of reliability as the ratio of true score variance to observed score variance, we can arrive at the following theorem:

$$\rho_{XX\prime} = \frac{\sigma_T^2}{\sigma_X^2} = \frac{\sigma_x^2 - \sigma_E^2}{\sigma_X^2} = 1 - \frac{\sigma_E^2}{\sigma_X^2}.$$

As the fraction of error variance to observed score variance tends to zero, the reliability then tends to 1. The CTT SEM, which assumes a homoscedastic error, is derived from the classical notion expressed above as $\sigma_X\sqrt{1 - \rho_{XX'}}$, where σ_X is the standard deviation of the scaled score, and $\rho_{XX'}$ is a reliability coefficient. Based on the definition of reliability, this formula can be derived as follows:

$$\rho_{XX'} = 1 - \frac{\sigma_E^2}{\sigma_X^2},$$

$$\frac{\sigma_E^2}{\sigma_X^2} = 1 - \rho_{XX'},$$

$$\sigma_E^2 = \sigma_X^2(1 - \rho_{XX'}), \text{ and}$$

$$\sigma_E = \sigma_X\sqrt{(1 - \rho_{XX'})}.$$

In general, the SEM is relatively constant across samples, as the group-dependent term, σ_X, can be shown to cancel out:

$$\sigma_E = \sigma_X\sqrt{(1 - \rho_{XX'})} = \sigma_X\sqrt{(1 - (1 - \frac{\sigma_E^2}{\sigma_X^2}))} = \sigma_X\sqrt{\frac{\sigma_E^2}{\sigma_X^2}} = \sigma_X \times \frac{\sigma_E}{\sigma_X} = \sigma_E.$$

This shows that the SEM in the CTT is assumed to be a homoscedastic error, irrespective of the standard deviation of a group.

By contrast, the SEMs in IRT vary over the ability continuum. These heterogeneous errors are a function of a test information function (TIF) that provides different information about examinees depending on their estimated abilities.

Because the TIF indicates the amount of information provided by the test at different points along the ability scale, its inverse indicates the lack of information at different points along the ability scale. This lack of information is the uncertainty, or the measurement error, of the score at various score points. Refer to Section 3, Reliability, for the derivation of heterogeneous measurement errors in IRT and a discussion of how these errors are aggregated over the score distribution to obtain a single, marginal, IRT-based reliability coefficient.

1.2 VALIDITY

The term *validity* refers to the degree to which "evidence and theory support the interpretations of test scores entailed by proposed uses of tests" (American Educational Research Association [AERA], American Psychological Association [APA], & National Council on Measurement in Education [NCME], 2014, p. 9). Messick (1989) defines validity as "an integrated evaluative judgment of the degree to which empirical evidence and theoretical rationales support the adequacy and appropriateness of inferences and actions based on test scores or other modes of assessment" (p. 13). Both definitions emphasize evidence and theory to support inferences and interpretations of test scores. The *Standards for Educational and Psychological Testing* (AERA, APA, & NCME, 2014) identify five sources of validity evidence that can be used in evaluating a proposed interpretation of test scores. When validating test scores, these sources of evidence should be carefully considered.

The first source of evidence for validity is the relationship between the test content and the intended test construct (see Section 4, Evidence of Content Validity). For test score inferences to support a validity claim, the items should be representative of the content domain, and the content domain should be relevant to the proposed interpretation of test scores. To determine content representativeness, diverse panels of content experts conduct alignment studies, in which experts review individual items and rate them based on how well they match the test specifications or cognitive skills required for a construct. Technology-enhanced items should be examined to ensure that no construct-irrelevant variance is introduced. If some aspect of the technology impedes or advantages a student in their responses to items, this could affect item responses and inferences regarding abilities on the measured construct. Refer to Volume 2, Test Development, of this technical report for details on the item development process.

The second source of validity evidence is based on "the fit between the construct and the detailed nature of performance or response actually engaged in by examinees" (AERA, APA, & NCME, 2014, p. 12). This evidence is collected by surveying examinees about their performance strategies or responses to specific items. Because items are developed to measure specific constructs and intellectual processes, evidence that examinees have engaged in relevant performance strategies to correctly answer the items supports the validity of the test scores.

The third source of evidence for validity is based on *internal structure*: the degree to which the relationships among test items and test components relate to the construct on which the proposed test scores are interpreted. Possible analyses to examine internal structure are dimensionality assessment, goodness-of-model-fit to data, and reliability analysis (see Section 3, Reliability; and Section 5, Evidence of Internal-External Structure, for details). It is also important to assess the degree to which the statistical relation between items and test components is invariant across groups. DIF analysis can be used to assess whether specific items function differently for subgroups of examinees (see Volume 1, Annual Technical Report).

The fourth source of evidence for validity is the relationship of test scores to external variables. The *Standards for Educational and Psychological Testing* (AERA, APA, & NCME, 2014) divides this source of evidence into three parts: (1) convergent and discriminant evidence; (2) test-criterion relationships; and (3) validity generalization. Convergent evidence supports the relationship between the test and other measures intended to assess similar constructs. Conversely, discriminant evidence delineates the test from other measures intended to assess different constructs. To analyze both convergent and discriminant evidence, a multitrait–multimethod matrix (MTMM) can be used. Additionally, test-criterion relationships indicate how accurately test scores predict criterion performance. The degree of accuracy depends mainly on the test's purpose, such as classification, diagnosis, or selection; test-criterion evidence is also used to investigate predictions of favoring different groups. Due to construct underrepresentation or construct-irrelevant components, the relation of test scores to a relevant criterion may differ from one group to another. Furthermore, validity generalization is related to whether the evidence is situation-specific or can be generalized across different settings and times. For example, sampling errors or range restriction may need to be considered in order to determine whether the conclusions of a test can be assumed for the larger population. Convergent and discriminant validity evidence are discussed in Section 5.2, Convergent and Discriminant Validity.

The fifth source of validity evidence is the intended and unintended consequences of test use, which should be included in the test-validation process. Determining the validity of the test should depend upon evidence directly related to the test; this process should not be influenced by external factors. For example, if an employer administers a test to determine hiring rates for different groups of people, an unequal distribution of skills related to the measurement construct does not necessarily imply a lack of validity for the test. However, if the unequal distribution of scores is in fact due to an unintended, confounding aspect of the test, this *would* interfere with the test's validity. As described in Volume 1, Annual Technical Report, and in this volume, test use should align with the intended purpose of the test.

Supporting a validity argument requires multiple sources of validity evidence. This enables one to evaluate whether sufficient evidence has been presented to support the intended uses and interpretations of the test scores. Thus, determining the validity of a test first requires an explicit statement regarding the intended uses of the test scores and, subsequently, evidence that the scores can be used to support these inferences.

2. Purpose of the Montana Science Assessment

The primary purpose of Montana's Summative Assessment System is to yield accurate information on students' performance of Montana's education standards. The Montana Science Assessment (MSA) measures the science knowledge and skills of Montana students in grades 5 and 8. The Montana Office of Public Instruction (OPI) provides an overview of the science assessment at https://opi.mt.gov/Educators/Teaching-Learning/K-12-Content-Standards/Science-Standards#88989724-science-assessment-updates. Information about the NGSS is available at www.nextgenscience.org.

The MSA supports instruction and student learning by measuring growth in student performance. Assessments can be used as indicators to determine whether students in Montana are ready with the knowledge and skills that are essential for college education and careers.

Montana's educational assessments also provide evidence for the requirements of state and federal accountability systems. Test scores can be employed to evaluate students' learning progress and to help teachers to improve their instruction, which in turn has a positive effect on students' learning over time.

The tests are constructed to measure student proficiency in accordance with best practice as described in the *Standards for Educational and Psychological Testing* (AERA, APA, & NCME, 2014). Item development adheres to the principles of universal design (UD) in order to ensure that all students have access to the test content. Volume 2, Test Development, describes in more detail the MSA's standards and test blueprints. Additional evidence of content validity can also be found in Section 4, Evidence of Content Validity. The MSA test scores are useful indicators for understanding individual students' academic performance of the Montana content standards and for evaluating whether students' performance is improving over time. Additionally, both individual and aggregated scores can be used for measuring reliability of the test. A discussion of test-score reliability can be found in Section 3, Reliability.

The MSA is a criterion-referenced test that is designed to measure student performance of the Montana Science Content Standards. As a comparison, norm-referenced tests are designed to rank or compare all students with one another. The MSA's standards and test blueprints are discussed in Volume 2, Test Development.

The scale score and relative strengths and weaknesses at the discipline level are provided for each student to indicate student strengths and weaknesses in different content areas of the test, relative to the other areas and to the district and state. These scores serve as useful feedback that teachers can use to tailor their instruction. To support the practical use of test scores across the state, we must examine their reliability coefficients and their validity.

3. RELIABILITY

Reliability indices based on classical test theory (CTT) are not appropriate for science assessments for two reasons. First, in spring 2022, the science test was administered under an adaptive test design. Potentially, each student received a unique set of items, whereas CTT-based reliability indices require that the same set of items be administered to a (large) group of students. Second, since item response theory (IRT) methods are used for calibration and scoring, the measurement error of ability estimates is not constant across the ability range, even for the same set of items. The reliability of science tests is computed as follows:

$$\bar{\rho} = [\sigma^2 - \left(\frac{\sum_{i=1}^{N} CSEM_i^2}{N}\right)]/\sigma^2,$$

where N is the number of students; $CSEM_i$ is the conditional standard errors of measurement (CSEM) of the overall ability estimate for student i; and σ^2 is the variance of the overall ability estimates. The higher the reliability coefficient, the greater the precision of the test.

The marginal reliability of science for the overall sample is reported by grade in Table 2. The overall reliability ranges from 0.88 to 0.89. Due to the new structure of the test, Cambium Assessment, Inc. (CAI) has also explored the relationships between reliability and other important factors, such as the effect of nuisance dimensions (see Section 5 of Volume 1, Annual Technical Report). It was found that if the local dependencies among assertions pertaining to the same item are ignored, the marginal reliability typically increases to 0.90 or above. Ignoring local dependencies can be achieved either by computing the maximum likelihood estimates (MLE) ability estimates under the unidimensional Rasch model or by setting the variance parameters to zero for all item clusters when computing the marginal maximum likelihood estimation (MMLE) ability estimates under the one-parameter logistic (1PL) bifactor model (see Section 6.1 of Volume 1, Annual Technical Report).

By ignoring the local dependencies, which are substantial for many item clusters, the reliability coefficient is overestimating the true reliability of the test. Note, however, that local dependencies are also present to some degree in traditional assessments that make use of item groups (e.g., a set of items relating to the same reading passage). Local dependencies are typically not accounted for by traditional assessments, and hence reported reliability coefficients may be overestimating to some degree the true reliability of these tests. The reliability coefficients are also reported for demographics subgroups in Appendix 4-A, Student Demographics and Reliability Coefficients.

Table 2. Marginal Reliability Coefficients

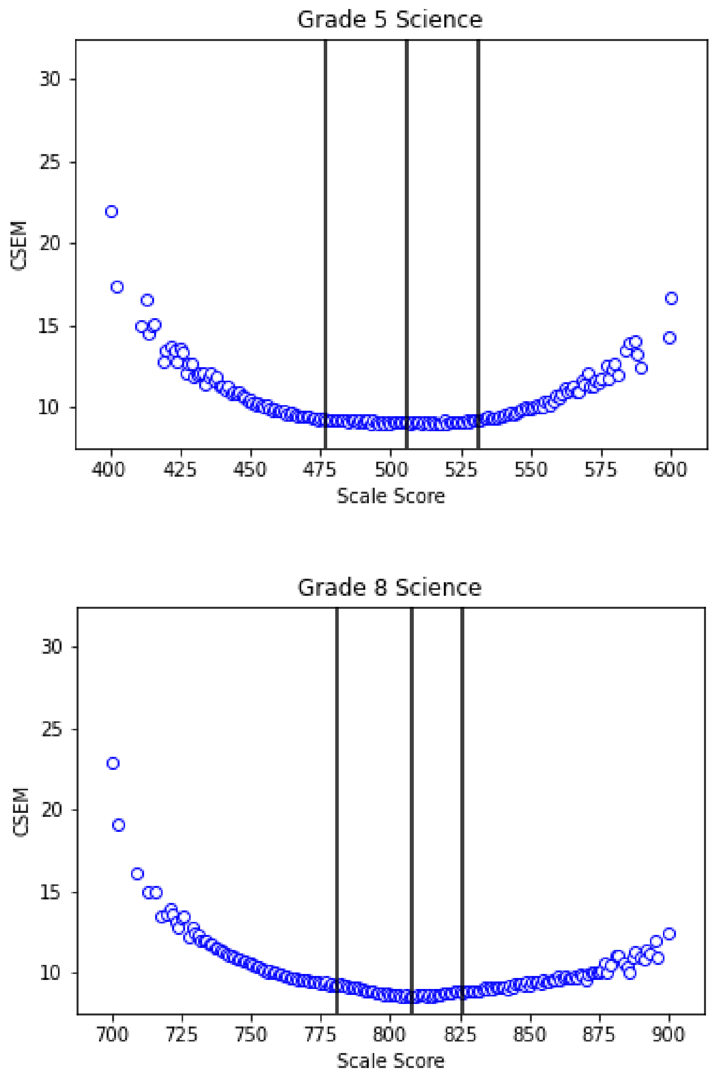
Grade	Sample Size	Reliability
5	11,068	0.88
8	11,159	0.89

3.1 STANDARD ERROR OF MEASUREMENT

The computation method of CSEM has been described in Section 6.4 of Volume 1, Annual Technical Report.

presents the average CSEM for each scale score. The lowest standard errors are observed near the proficiency cut (the middle vertical line) for all grades, which is a desirable test property. The CSEM at each scale score is reported in Appendix 4-B, Conditional Standard Error of Measurement.

Figure 1. Conditional Standard Errors of Measurement



3.2 Reliability of Performance Classification

When student performance is reported in terms of performance levels, the reliability of classifying students into a specific level can be computed in terms of the likelihood of accurate and consistent classification, as specified in Standard 2.16 of the *Standards for Educational and Psychological Testing* (AERA, APA, & NCME, 2014).

The reliability of performance classification can be examined in terms of *classification accuracy* (CA) and *classification consistency* (CC). CA refers to the agreement between the classifications based on the form taken and the classifications that would be made based on the students' true scores if hypothetically they could be obtained. CC refers to the agreement between the classifications based on the form taken and the classifications that would be made based on an alternate, equivalently constructed test form.

In reality, the true ability is unknown, and students are not administered an alternate, equivalent form. Therefore, CA and CC are estimated on the basis of students' item scores, the item parameters, and the assumed latent ability distribution as described in the following sections. The true score is an expected value of the test score with measurement error.

For student j, the student's estimated ability is $\hat{\theta}_j$ with a standard error of measurement (SEM) of $se\left(\hat{\theta}_j\right)$; and the estimated ability is distributed as $\hat{\theta}_j \sim N\left(\theta_j, se^2\left(\hat{\theta}_j\right)\right)$, assuming a normal distribution, where θ_j is the unknown true ability of student j. The probability of the true score at performance level l ($l = 1, \cdots, L$) is estimated as

$$p_{jl} = p(c_{Ll} \leq \theta_i < c_{Ul}) = p\left(\frac{c_{Ll}-\hat{\theta}_j}{se(\hat{\theta}_j)} \leq \frac{\theta_j-\hat{\theta}_j}{se(\hat{\theta}_j)} < \frac{c_{Ul}-\hat{\theta}_j}{se(\hat{\theta}_j)}\right) = p\left(\frac{\hat{\theta}_j-c_{Ul}}{se(\hat{\theta}_j)} < \frac{\hat{\theta}_j-\theta_j}{se(\hat{\theta}_j)} \leq \frac{\hat{\theta}_j-c_{Ll}}{se(\hat{\theta}_j)}\right) = \Phi\left(\frac{\hat{\theta}_j-c_{Ll}}{se(\hat{\theta}_j)}\right) - \Phi\left(\frac{\hat{\theta}_j-c_{Ul}}{se(\hat{\theta}_j)}\right),$$

where c_{Ll} and c_{Ul} denote the score corresponding to the lower and upper limits of performance level l, respectively.

3.2.1 Classification Accuracy

Using p_{jl}, an $L \times L$ matrix $\boldsymbol{E_A}$ can be calculated. Each element E_{Akl} of matrix $\boldsymbol{E_A}$ represents the expected number of students to score at level l (based on their true scores) given students from observed level k, and can be calculated as

$$E_{Akl} = \sum_{pl_j \in k} p_{jl},$$

where pl_j is the jth student's observed performance level. The classification accuracy (*CA*) at level l is estimated as

$$CA_l = \frac{E_{Akl}}{N_k},$$

where N_k is the observed number of students scoring in performance level k.

The CA for the pth cut (CAC) is estimated by forming square partitioned blocks of the matrix $\boldsymbol{E_A}$ and taking the summation over all elements within the block as follows:

$$CAC = \left(\sum_{k=1}^{p}\sum_{l=1}^{p} E_{Akl} + \sum_{k=p+1}^{L}\sum_{l=p+1}^{L} E_{Akl}\right)/N,$$

where N is the total number of students.

The overall CA is estimated from the diagonal elements of the matrix:

$$CA = \frac{tr(\boldsymbol{E_A})}{N}.$$

Table 3 provides the overall CA and the CA for the individual cuts. The overall CA of the test ranges from 76.09% to 76.98%. The individual cut accuracy rates are high across all grades and forms, with the minimum value being 88.95% for grade 8. This denotes that we can accurately differentiate students between adjacent performance levels more than 88% of the time in the spring 2023 Montana Science Assessment (MSA). The CA for demographic subgroups is presented in Appendix 4-C, Classification Accuracy and Consistency Indices by Subgroups.

Table 3. Spring 2023 Classification Accuracy Index

Grade	Overall Accuracy (%)	Cut Accuracy (%)		
		Level 2 Cut	*Level 3 Cut*	*Level 4 Cut*
5	76.98	92.55	90.18	94.22
8	76.09	93.44	88.95	93.53

3.2.2 Classification Consistency

Assuming the test is administered twice independently to the same group of students, similarly to accuracy, a $L \times L$ matrix $\boldsymbol{E_C}$ can be constructed. The element of $\boldsymbol{E_C}$ is populated by

$$E_{Ckl} = \sum_{j=1}^{N} p_{jl} p_{jk},$$

where p_{jl} is the probability of the true score at performance level l in test one, and p_{jk} is the probability of the true score at performance level k in test two for the jth student. The classification consistency index for the cuts (*CCC*) and overall classification consistency (*CC*) were estimated in a way similar to CAC and CA.

$$CCC = \left(\sum_{k=1}^{p} \sum_{l=1}^{p} E_{Ckl} + \sum_{k=p+1}^{L} \sum_{l=p+1}^{L} E_{Ckl}\right)/N,$$

and

$$CC = \frac{tr(\boldsymbol{E_C})}{N}.$$

Table 4 provides the overall CC and the CC for the cuts. CC rates can be lower than CA; the consistency is based on two tests with measurement errors, but the accuracy is based on one test with a measurement error and the true score. The accuracy and consistency rates for each performance level are higher for the levels with a smaller standard error. The CC for demographic subgroups is presented in Appendix 4-C, Classification Accuracy and Consistency Indices by Subgroups.

Table 4. Spring 2023 Classification Consistency Index

Grade	Overall Consistency (%)	Cut Consistency (%)		
		Level 2 Cut	*Level 3 Cut*	*Level 4 Cut*
5	68.00	89.43	86.28	91.85
8	67.39	90.66	84.69	90.83

3.3 Precision at Cut Scores

Table 5 presents the mean CSEM at each performance level by grade. The table also includes performance level cut scores and associated CSEMs. The CSEM at each scale score is reported in Appendix 4-B, Conditional Standard Error of Measurement.

Table 5. Spring 2023 Performance Levels and Associated Conditional Standard Errors of Measurement

Grade	Performance Level	Mean CSEM	Cut Score (Scale Score)	CSEM at Cut Score
5	1	9.92	-	-
	2	9.14	477	9.23
	3	9.10	506	9.11
	4	9.72	531	9.21
8	1	9.98	-	-
	2	8.82	781	9.23
	3	8.65	808	8.57
	4	9.17	826	8.81

4. Evidence of Content Validity

This section demonstrates that the knowledge and skills assessed by the Montana Science Assessment (MSA) are representative of the content standards of the larger knowledge domain. We describe the content standards for the MSA and discuss the test development process and mapping of MSA tests to the *Standards for Educational and Psychological Testing* (AERA, APA, & NCME, 2014). A complete description of the test development process can be found in Volume 2, Test Development.

4.1 Content Standards

The Montana Science Content Standards were adopted by the Montana Office of Public Instruction (OPI) in 2016. The standards are available for review at the following URL: https://opi.mt.gov/Educators/Teaching-Learning/K-12-Content-Standards/Science-

Standards#88989724-science-assessment-updates. Blueprints were developed to ensure that the tests and the items were aligned to the standards that they were intended to measure. A complete description of the blueprint and test construction process can be found in Volume 2, Test Development.

Table 6 presents the disciplines by grade, as well as the number of operational items administered that measured each discipline.

Table 6. Number of Spring 2023 Items for Each Discipline

Grade	Reporting Category	Item Cluster	Stand-Alone Item
5	Earth and Space Sciences	32	42
	Life Sciences	23	37
	Physical Sciences	32	47
8	Earth and Space Sciences	22	35
	Life Sciences	21	39
	Physical Sciences	29	38

5. Evidence of Internal-External Structure

In this section, the internal structure of the assessment is explored using the scores provided at the discipline level. The relationship between the discipline scores is just one indicator of the test dimensionality. The Montana Science Assessment (MSA) is calibrated with the Rasch testlet model (Wang & Wilson, 2005). The testlet model is a high-dimensional model that incorporates a nuisance dimension for each item cluster (and stand-alone items with four or more assertions) in addition to an overall dimension representing overall proficiency. This innovative approach differs from the traditional one, in which local dependencies are ignored. Validity evidence for the internal structure will focus on the presence of cluster effects and how substantial they are. Confirmatory factor analysis (CFA) is also used to evaluate the fit of the IRT model and to compare the model with alternative models, including those with a simpler internal structure (i.e., unidimensional models without cluster effects) and models with a more elaborate internal structure (refer to Section 5.4, Confirmatory Factor Analysis).

Another pathway is to explore observed correlations between the discipline scores. However, as each discipline is measured with a small number of items, the standard errors of the observed scores within each discipline are typically larger than the standard error of the total test score. Disattenuating for measurement error could offer some insight into the theoretical true score correlations. Both observed correlations and disattenuated correlations are provided in the following section.

5.1 Correlations Among Discipline Scores

Table 7 presents the observed and disattenuated correlation matrix of the discipline scores. The observed correlations ranged from 0.66 to 0.72, and the disattenuated correlations ranged from 0.94 to 0.99.

In some instances, the observed correlations were lower than expected. However, as previously noted, the correlations were subject to a large amount of measurement error at the discipline level due to the limited number of items from which the scores were derived. Consequently, interpretation of these correlations, as either high or low, should be made cautiously. After correcting for measurement error, the correlations between the discipline scores become very high. The disattenuated correlations were close to 1, supporting the use of a psychometric model that did not include a separate dimension for each of the three disciplines.

Table 7. Spring 2023 Correlations Among Disciplines

Grade	Reporting Category	Earth and Space Sciences (ESS)	Life Sciences (LS)	Physical Sciences (PS)
5	ESS	0.75*	0.96	0.96
	LS	0.69	0.70*	0.94
	PS	0.70	0.66	0.70*
8	ESS	0.73*	0.99	0.96
	LS	0.72	0.73*	0.98
	PS	0.70	0.71	0.73*

*Diagonal value represents marginal reliability for each discipline. Observed correlations are below the diagonal, and disattenuated are above.

5.2 Convergent and Discriminant Validity

Collectively, Standard 1.16 through Standard 1.19 of the *Standards for Educational and Psychological Testing* (AERA, APA, & NCME, 2014) emphasize practices that provide evidence of convergent and discriminant validity. Part of validity evidence involves demonstrating that assessment scores are related as expected with criteria and other variables for all student groups. However, a second, independent test measuring the same science construct as the MSA, which could easily permit for a cross-test set of correlations, was not available. Alternatively, the correlations between subscores were examined. The *a priori* expectation is that subscores within the same subject (e.g., correlation of science disciplines within science) will correlate more positively than subscores across subjects (e.g., correlation of science disciplines with reporting categories within mathematics). These correlations are based on a small number of items; consequently, the observed score correlations will be smaller in magnitude as a result of the larger measurement error at the subscore level. For this reason, both the observed score and the disattenuated correlations are provided.

Observed and disattenuated subscore correlations were calculated both within and across subjects. The pattern was generally consistent with the *a priori* expectation that subscores within a test correlate higher than correlations between tests measuring a different construct. The correlations between reporting categories from science, ELA, and mathematics are presented in Table 8 and Table 9. On the diagonal, the reliability coefficient of the reporting category is shown.

Table 8. Correlations Across Spring 2023 Subjects, Grade 5

Subject	Number of Students	Reporting Category	Science			English Language Arts (ELA)			Mathematics		
			ESS	*LS*	*PS*	*R*	*WR*	*L*	*CP*	*PS*	*CR*
Science	10,991	Earth and Space Sciences (ESS)	0.75*	0.96	0.96	0.88	0.88	0.89	0.81	0.92	0.88
		Life Sciences (LS)	0.69	0.69*	0.94	0.89	0.87	0.89	0.80	0.90	0.87
		Physical Sciences (PS)	0.69	0.66	0.70*	0.88	0.87	0.89	0.82	0.90	0.88
ELA		Reading (R)	0.67	0.65	0.65	0.77*	0.97	0.97	0.79	0.90	0.86
		Writing and Research/Inquiry (WR)	0.69	0.66	0.66	0.77	0.82*	0.96	0.81	0.92	0.88
		Listening (L)	0.59	0.57	0.57	0.65	0.67	0.59*	0.81	0.92	0.88
Mathematics		Concepts and Procedures (CP)	0.66	0.62	0.64	0.65	0.69	0.58	0.88*	1.00	0.98
		Problem Solving, Modeling, and Data Analysis (PS)	0.62	0.59	0.59	0.62	0.65	0.55	0.73	0.61*	1.00
		Communicating and Reasoning (CR)	0.62	0.59	0.60	0.62	0.65	0.55	0.75	0.68	0.67*

*Diagonal value represents the reliability coefficient of the reporting category. Observed correlations are below the diagonal, and disattenuated are above. Disattenuated correlations larger than 1 were truncated to 1.

Table 9. Correlations Across Spring 2023 Subjects, Grade 8

Subject	Number of Students	Reporting Category	Science			English Language Arts (ELA)			Mathematics		
			ESS	*LS*	*PS*	*R*	*WR*	*L*	*CP*	*PS*	*CR*
Science	11,005	Earth and Space Sciences (ESS)	0.73*	0.99	0.96	0.85	0.84	0.86	0.85	0.95	0.89
		Life Sciences (LS)	0.72	0.72*	0.97	0.86	0.85	0.87	0.85	0.95	0.91
		Physical Sciences (PS)	0.70	0.71	0.73*	0.83	0.82	0.84	0.85	0.94	0.88
ELA		Reading (R)	0.63	0.64	0.61	0.75*	0.99	0.99	0.83	0.94	0.89
		Writing and Research/Inquiry (WR)	0.63	0.63	0.61	0.75	0.76*	0.98	0.83	0.93	0.88
		Listening (L)	0.56	0.56	0.54	0.65	0.65	0.58*	0.84	0.94	0.89
Mathematics		Concepts and Procedures (CP)	0.67	0.67	0.67	0.67	0.67	0.59	0.86*	1.00	1.00
		Problem Solving, Modeling, and Data Analysis (PS)	0.61	0.61	0.61	0.62	0.62	0.54	0.75	0.57*	1.00
		Communicating and Reasoning (CR)	0.60	0.60	0.59	0.60	0.60	0.53	0.74	0.68	0.61*

*Diagonal value represents the reliability coefficient of the reporting category. Observed correlations are below the diagonal, and disattenuated are above. Disattenuated correlations larger than 1 were truncated to 1.

Additionally, the correlation was computed among the overall scores for the three tested subjects: ELA, mathematics, and science. Correlations are presented in Table 10 and are relatively high, between 0.76 and 0.81.

Table 10. Correlations Across Spring 2023 ELA, Mathematics, and Science Scores

Grade	*N*	English Language Arts (ELA) and Mathematics	ELA and Science	Mathematics and Science
5	10,991	0.77	0.81	0.77
8	11,005	0.76	0.76	0.78

5.3 CLUSTER EFFECTS

The MSA is calibrated with the Rasch testlet model (Wang & Wilson, 2005). The testlet model is a high-dimensional model that incorporates a nuisance dimension for each item cluster in addition to an overall dimension representing overall proficiency. Section 5.1 of Volume 1, Annual Technical Report, presents a detailed description of the IRT model. The internal (latent) structure of the model is presented in Figure 6. The innovative psychometric approach for the assessment differs from the traditional approach, in which local dependencies are ignored. The validity evidence for the internal structure presented in this section relates to the presence of cluster effects (i.e., nuisance dimensions) and how substantial they are.

Simulation studies conducted by Rijmen, Jiang, and Turhan (2018) confirmed that both the item difficulty parameters and the cluster variances were recovered well for the Rasch testlet model under a variety of conditions. Cluster effects with a range of magnitudes were recovered well. The results obtained by Rijmen et al. (2018) confirmed earlier findings reported in the literature (e.g., Bradlow, Wainer, & Wang, 1999) under conditions that were chosen to closely resemble the assessment. For example, in one of the studies, the item location parameters and cluster variances used to simulate data were based on the results of a pilot study.

We examined the distribution of cluster variances obtained from the 2022 IRT calibrations for the entire bank used across all states that participate in the Memorandum of Understanding (MOU) item-sharing agreement and the states that rely on the science ICCR item pool.

For elementary school, the estimated value of the cluster variances of all operational, scored items ranged from 0 to 7.72, with a median value of 0.43 and a mean value of 0.64. As a comparison, the estimated variance parameter of the overall dimension for Montana elementary school in 2022 was $\hat{\sigma}^2_{\theta MT}$=0.87.

For middle school, the estimated value of the cluster variances of all operational, scored items ranged from 0 to 4.68, with a median value of 0.43 and a mean value of 0.65. The estimated variance parameter of the overall dimension for Montana middle school in 2022 was $\hat{\sigma}^2_{\theta MT}$= 0.84.

Figure 2 and Figure 3 present histograms of the cluster variances expressed as the proportion of the systematic variance due to the cluster variance for each cluster (computed as $\eta_g = \frac{\hat{\sigma}^2_g}{\hat{\sigma}^2_{\theta_{MT}} + \hat{\sigma}^2_g}$), where $\hat{\sigma}^2_{\theta_{MT}}$is the variance estimate of the overall proficiency of Montana students. The variance proportion shows the relative magnitude of the variance of a cluster compared to the variance of the overall dimension. For instance, if the variance proportion of a cluster is larger than 0.5, then the cluster variance is larger than the overall variance; otherwise, the cluster variance is smaller than the overall variance. For both grade bands, a wide range of cluster variances is observed. These results indicate that, for all grades, cluster effects can be substantial and provide evidence for the appropriateness of a psychometric model that explicitly takes local dependencies among the assertions of an item cluster into account.

Figure 2. Cluster Variance Proportion for Operational Items in Elementary School

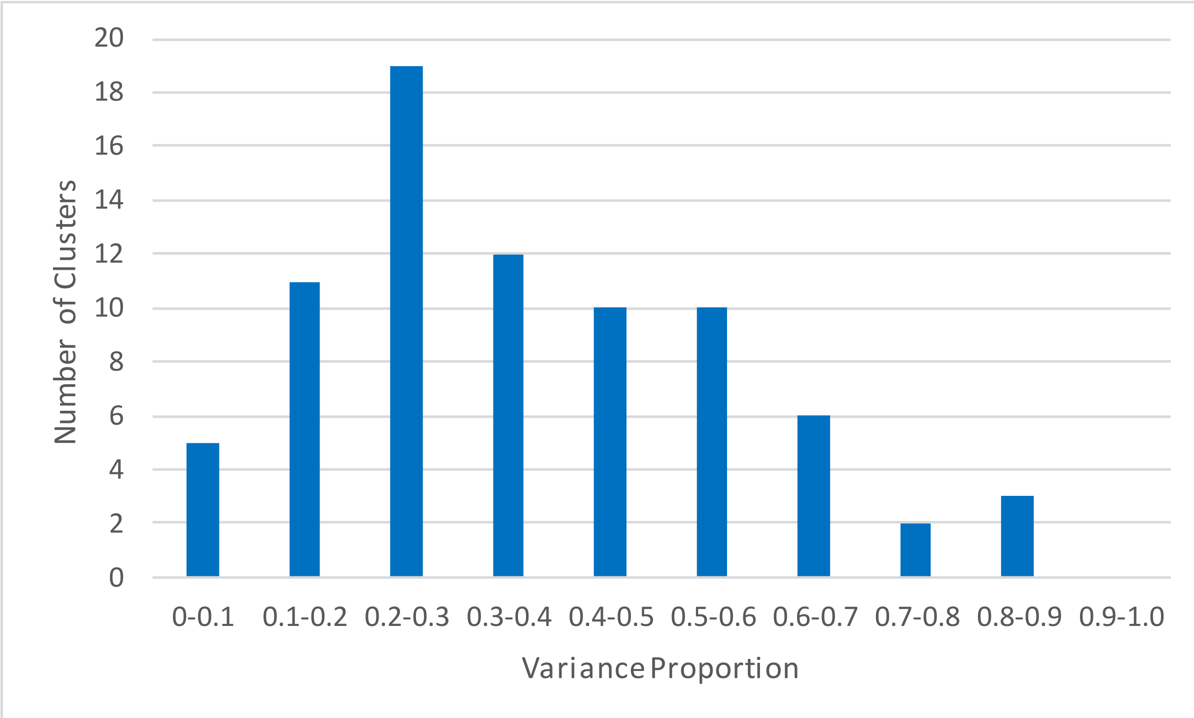
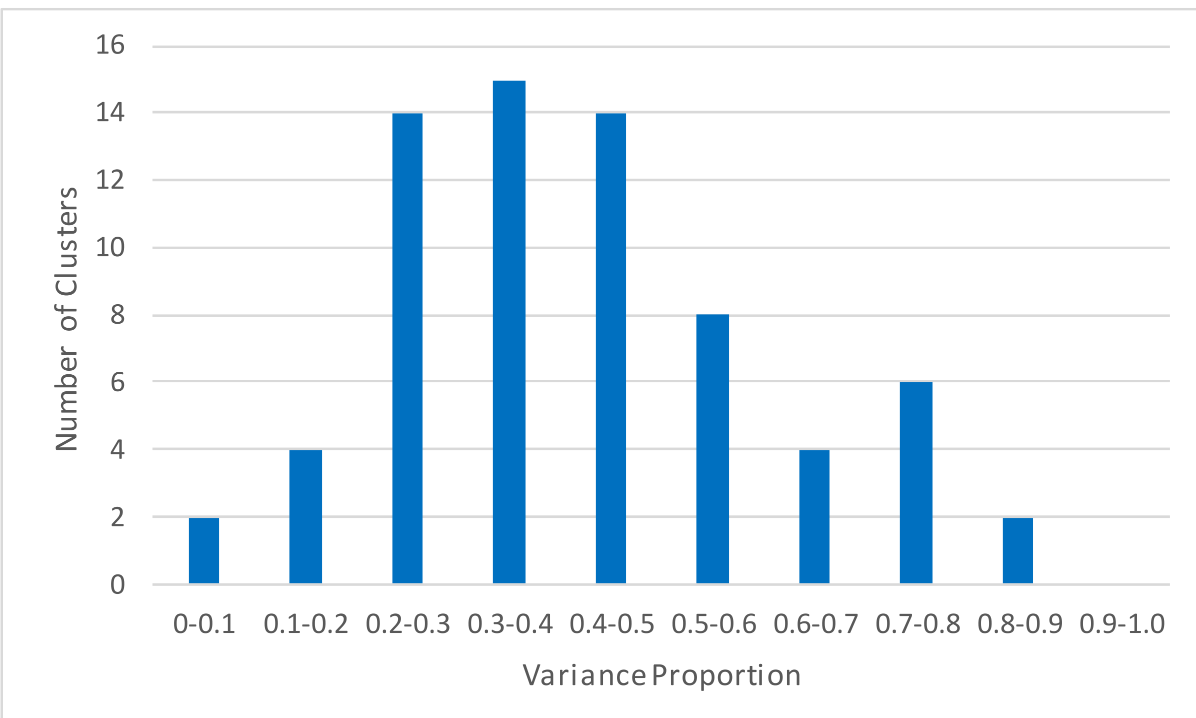


Figure 3. Cluster Variance Proportion for Operational Items in Middle School



5.4 CONFIRMATORY FACTOR ANALYSIS

In Section 5.3, Cluster Effects, evidence is presented for the existence of substantial cluster effects. In this section, the internal structure of the IRT model used for calibrating the item parameters is further evaluated using CFA. In addition, alternative models are considered, including models with a simpler internal structure (e.g., unidimensional models) and models with a more elaborate internal structure.

Estimation methods for CFA for discrete observed variables are not well suited for incomplete data collection designs where each case has data only on a subset of the set of observed variables. The linear-on-the-fly testing (LOFT) design results in sparse data matrices because every student is responding to only a small number of items relative to the size of the item pool; so data are missing on most of the manifest variables for any given student. In 2018 and 2019, a LOFT design was used for all operational science assessments inspired by the NGSS framework, except in Utah. As a result, the student responses in these other states were not readily amenable to the application of CFA techniques.

The 2018 Utah operational field test for science made use of a set of fixed-form tests for each grade. Therefore, the data for each fixed-form test were complete, and the fixed-form tests were amenable to CFA. Although the Utah science standards were grade-specific for middle school, they were developed under a framework similar to the one developed for the NGSS; and a crosswalk was available between both sets of standards.

Utah is part of the MOU, and many of the other states that take part in the MOU also use the middle school items developed for and owned by Utah. Taken together, analyzing the fixed science forms that were administered in Utah in 2018 can provide evidence with respect to the internal structure of the MSA.

In 2018, Utah’s science assessments comprised a set of fixed-form tests per grade, and all items in these forms were clusters. The number of fixed-form tests varied by grade, but within each grade the total number of clusters was the same across forms. However, some items were rejected during the rubric validation or data review and were removed from this analysis. All students with a “completed” status were included in the CFA. The percentage of students per grade that had a status other than “completed” was less than 0.85%. Table 11 summarizes the number of forms included in this analysis, the number of clusters per discipline (range across forms), the number of assertions (range across forms), and the number of students (range across forms) for each of the grades.

Table 11. Numbers of Forms, Clusters per Discipline (Range Across Forms), Assertions per Form (Range Across Forms), and Students per Form (Range Across Forms)

Grade	Number of Fixed Forms	Number of Clusters per Discipline in Each Form			Number of Assertions per Form	Number of Students per Form
		Physical Sciences	*Earth and Space Sciences*	*Life Sciences*		
6	3	2	2–3	2–3	74–83	6,804–6,881
7	6	2	2	5	83–89	3,822–3,890

Grade	Number of Fixed Forms	Number of Clusters per Discipline in Each Form			Number of Assertions per Form	Number of Students per Form
		Physical Sciences	*Earth and Space Sciences*	*Life Sciences*		
8	3	6–7	2	2	93–100	5,061–5,104

The factor structure of a testlet model, which is the model used for calibration, is formally equivalent to a second-order model. Specifically, the testlet model is the model obtained after a Schmid–Leiman transformation of the second-order model (Li, Bolt, & Fu, 2006; Rijmen, 2009; Yung, Thissen, & McLeod, 1999). In the corresponding second-order model, the group of assertions related to a cluster are indicators of the cluster, and each cluster is an indicator of overall science performance. Because assertions are not pure indicators of a specific factor, each assertion has a corresponding error component. Similarly, clusters include an error component indicating they are not pure indicators of the overall science performance.

CAI used CFA to evaluate the fit of the second-order model described previously to student data from spring 2018. Three additional structural models were included in the analysis as well. In the first model, only one factor represented overall science performance. All assertions are indicators of this overall proficiency factor. The first model was a testlet model where all cluster variances were zero. In the second model, assertions were indicators of the corresponding science discipline, and each discipline was an indicator of the overall science performance. This was a second-order model with science disciplines rather than clusters as first-order factors. This model did not take the cluster effects into account. In the last, most general model, assertions were indicators of the corresponding cluster, and clusters were indicators of the corresponding science discipline, with disciplines being indicators of the overall science performance.

For the sake of simplicity, the models in the analysis are here referred to as

- Model 1—Assertions-Overall Science (one-factor model)
- Model 2—Assertions-Disciplines-Overall Science (second-order model)
- Model 3—Assertions-Clusters-Overall Science (second-order model)
- Model 4—Assertions-Clusters-Disciplines-Overall Science (third-order model)

Figure 4 through Figure 7 illustrate these four structural models. Model 1 is nested within Models 2, 3, and 4. Also, Models 2 and 3 are nested within Model 4. The paths from the factors to the assertions represent the first-order factor loadings. Note that all four models include factor loadings for the assertions, which differs from the calibration model where all the discrimination parameters of the assertions were set to 1.

Figure 4. One-Factor Structural Model (Assertions-Overall): "Model 1"

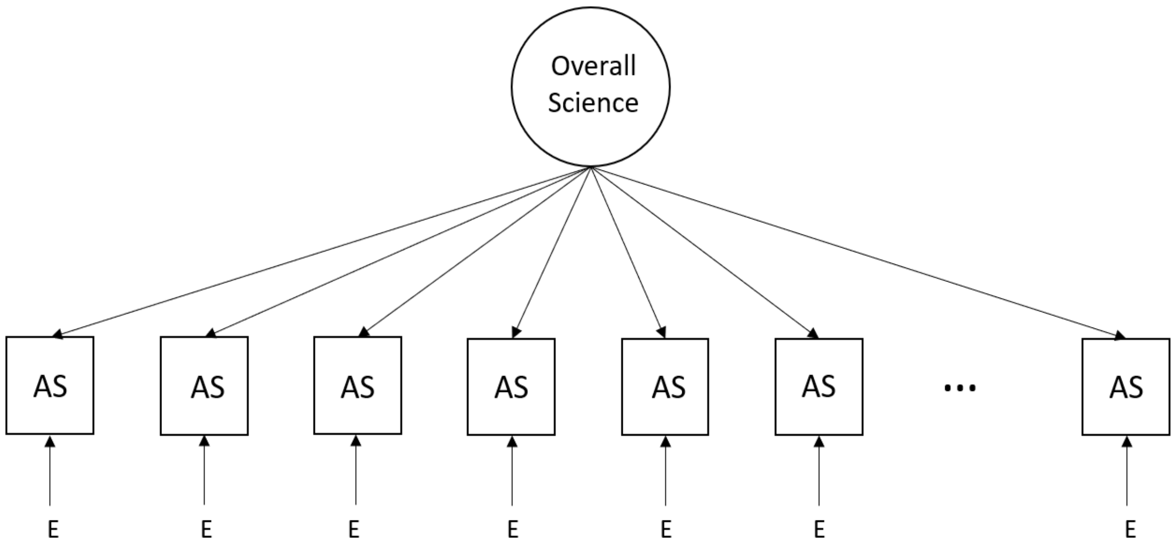


Figure 5. Second-Order Structural Model (Assertions-Disciplines-Overall): "Model 2"

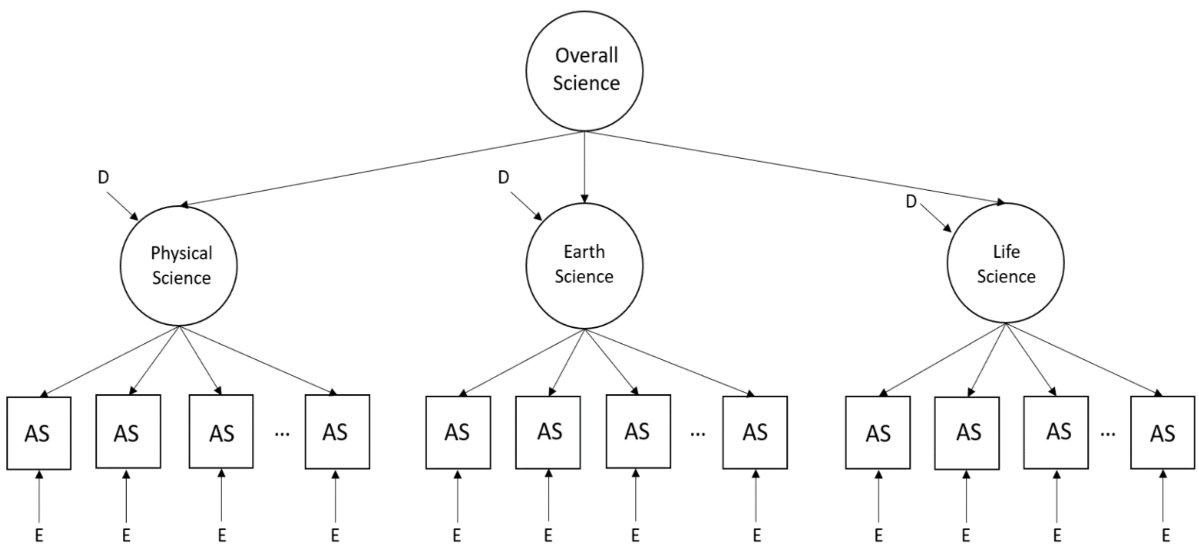


Figure 6. Second-Order Structural Model (Assertions-Clusters-Overall): "Model 3"

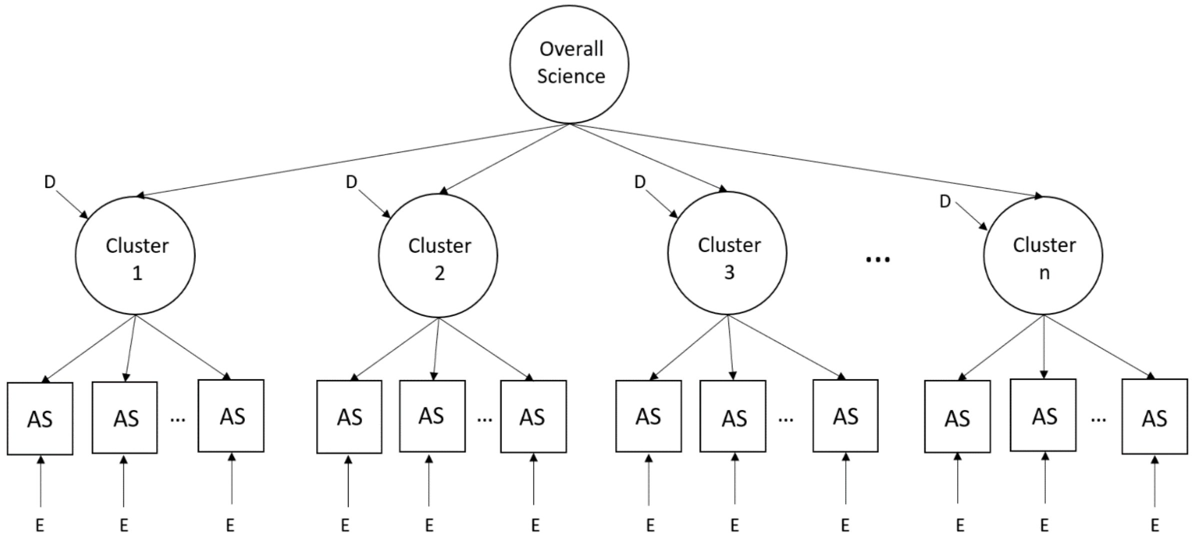
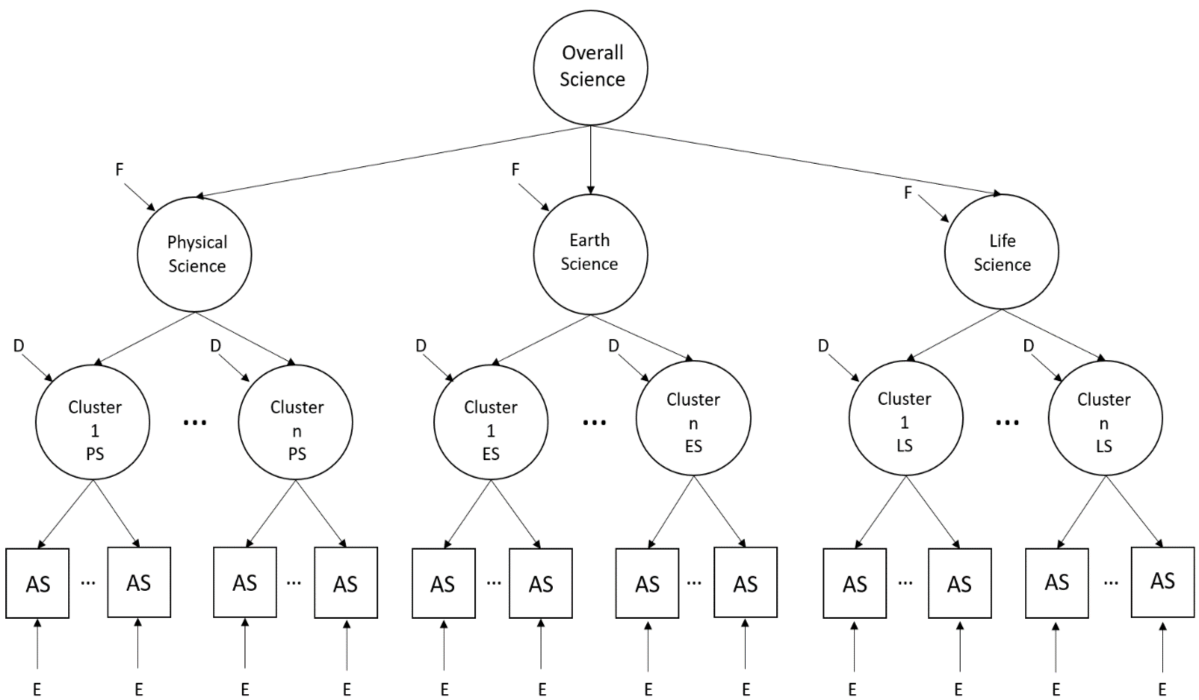


Figure 7. Third-Order Structural Model (Assertions-Clusters-Disciplines-Overall): "Model 4"



5.4.1 Results

For each test form, fit measures were computed for each of the four models. The fit measures used to evaluate goodness-of-fit were the comparative fit index (CFI), the Tucker-Lewis index (TLI), the root mean square error of approximation (RMSEA), and the standardized root mean squared residual (SRMR). CFI and TLI are relative fit indices, meaning they evaluate model fit by comparing the model of interest to a baseline model. RMSEA and SRMR are indices of absolute fit. Table 12 provides a list of these measures along with the corresponding thresholds indicating a good fit.

Table 12. Guidelines for Evaluating Goodness of Fit

Goodness-of-Fit Measure*	Indication of Good Fit
CFI	≥ 0.95
TLI	≥ 0.95
RMSEA	≤ 0.06
SRMR	≤ 0.08

*Brown, 2015; Hu & Bentler, 1999

Table 13 through Table 15 show the goodness-of-fit statistics for grades 6 through 8, respectively.[1] Numbers in bold indicate those indices that did not meet the criteria established in Table 12. Across all grades and models, the following conclusions can be drawn:

- Model 1 shows the most misfit across grades and forms.
- Across forms, Model 3 generally shows more improvement in model fit relative to Model 1 than Model 2 does (i.e., higher values for CFI and TLI and lower values for RMSEA and SRMR). This means that accounting for the clusters resulted in a higher improvement in model fit over a single factor model than accounting for disciplines.
- Model 4 does not show improvement in model fit over Model 3. Fit measures remained the same (or had a difference of 0.001 or smaller in very few cases) across forms for Models 3 and 4. Hence, including the disciplines in the model (when clusters were accounted for) did not improve model fit.
- Overall model fit for Models 3 and 4 decreases with decreasing grades. For grade 8, all fit indices for Models 3 and 4 indicate good model fit for all three forms. For grade 7, all fit indices for Models 3 and 4 indicate good fit for two out of the six forms, and the degree of misfit for the other four forms is small. For grade 6, all three forms have fit indices above the threshold values for at least one of the absolute fit indices for Models 3 and 4. The amount of misfit is small for the RMSEA but more substantial for the SRMR for two out of the three forms.

[1] For very few assertions per form and models, some error variances were slightly below 0. For grade 6, 1–2 assertions per form and model had error variance below 0, with the lowest error variance being −0.027. For grade 7, Forms 1, 2, 5, and 6 had a single negative error variance for 1 assertion in Models 3 and 4, with the lowest error variance being −0.099. Form 4 had 1–2 assertions with negative error variance in each model, and the lowest error variance was −0.102. For grade 8, there were no assertions with negative error variances for any of the forms and models.

Table 13. Fit Measures per Model and Form, Grade 6

Model	Form	CFI	TLI	RMSEA	SRMR
Model 1 Assertions-Overall (one-factor model)	**1**	0.995	0.995	**0.106**	**0.163**
	2	0.997	0.997	**0.093**	**0.148**
	3	0.995	0.995	**0.109**	**0.161**
Model 2 Assertions-Disciplines-Overall (second-order model)	**1**	0.996	0.996	**0.089**	**0.144**
	2	0.998	0.998	**0.078**	**0.128**
	3	0.997	0.997	**0.087**	**0.135**
Model 3 Assertions-Clusters-Overall (second-order model)	**1**	0.998	0.998	**0.065**	**0.107**
	2	0.999	0.999	0.056	**0.095**
	3	0.998	0.998	**0.067**	**0.104**
Model 4 Assertions-Clusters-Disciplines-Overall (third-order model)	**1**	0.998	0.998	**0.065**	**0.107**
	2	0.999	0.999	0.056	**0.095**
	3	0.998	0.998	**0.067**	**0.104**

Note. Numbers in bold do not meet the criteria for goodness of fit.

Table 14. Fit Measures per Model and Form, Grade 7

Model	Form	CFI	TLI	RMSEA	SRMR
Model 1 Assertions-Overall (one-factor model)	**1**	**0.892**	**0.889**	0.060	0.074
	2	**0.938**	**0.936**	**0.083**	**0.109**
	3	**0.940**	**0.939**	0.052	0.065
	4	**0.937**	**0.936**	**0.068**	**0.114**
	5	**0.939**	**0.937**	**0.093**	**0.119**
	6	**0.898**	**0.895**	0.056	0.071
Model 2 Assertions-Disciplines-Overall (second-order model)	**1**	**0.908**	**0.906**	0.055	0.073
	2	0.962	0.961	**0.065**	**0.088**
	3	0.950	**0.949**	0.048	0.063
	4	0.955	0.954	0.058	**0.094**
	5	0.959	0.957	**0.077**	**0.103**
	6	**0.906**	**0.903**	0.054	0.070
Model 3 Assertions-Clusters-Overall (second-order model)	**1**	**0.938**	**0.937**	0.046	0.072
	2	0.974	0.973	0.054	**0.082**
	3	0.967	0.966	0.039	0.055
	4	0.977	0.976	0.041	0.072
	5	0.975	0.974	0.060	**0.089**
	6	**0.932**	**0.930**	0.046	0.072

Model	Form	CFI	TLI	RMSEA	SRMR
Model 4 Assertions-Clusters-Disciplines-Overall (third-order model)	**1**	**0.939**	**0.937**	0.045	0.072
	2	0.974	0.973	0.054	**0.082**
	3	0.967	0.966	0.039	0.055
	4	0.977	0.976	0.041	0.072
	5	0.975	0.974	0.060	**0.089**
	6	**0.932**	**0.930**	0.046	0.072

Note. Numbers in bold do not meet the criteria for goodness of fit.

Table 15. Fit Measures per Model and Form, Grade 8

Model	Form	CFI	TLI	RMSEA	SRMR
Model 1 Assertions-Overall (one-factor model)	**1**	**0.929**	**0.927**	0.043	0.060
	2	0.959	0.958	0.042	0.056
	3	**0.943**	**0.941**	0.052	0.074
Model 2 Assertions-Disciplines - Overall (second-order model)	**1**	**0.934**	**0.932**	0.041	0.060
	2	0.963	0.963	0.040	0.056
	3	0.950	**0.949**	0.049	0.072
Model 3 Assertions-Clusters-Overall (second-order model)	**1**	0.953	0.952	0.034	0.057
	2	0.974	0.973	0.034	0.054
	3	0.970	0.969	0.038	0.064
Model 4 Assertions-Clusters-Disciplines-Overall (third-order model)	**1**	0.953	0.952	0.034	0.057
	2	0.974	0.974	0.033	0.053
	3	0.970	0.969	0.038	0.064

Note. Numbers in bold do not meet the criteria for goodness of fit.

For Models 3 and 4, grade 6 showed some degree of misfit across all three forms according to the measures of absolute model fit, especially for the SRMR. Further examination indicated that the lack of fit could be attributed to a single item that was common to all three grade 6 forms that were part of this factor analysis study. Once this item was removed, only two forms had two or more clusters per discipline. The fit for both forms improved drastically in Models 3 and 4, with all fit measures except the SRMR for one form meeting the criteria for model fit. The SRMR value that exceeded the threshold value did so barely, with a value of 0.083. Table 16 shows the fit measures for grade 6 after removal of the item causing misfit. Note that, unlike Models 3 and 4, Models 1 and 2 still did not meet the criteria of model fit after removing the item.[2]

[2] One assertion per model in form 1 and one assertion on three of the models in form 2 had error variance below 0, with the lowest error variance being −0.027.

Table 16. Fit Measures per Model and Form, Grade 6, with One Cluster Removed

Model	Form	CFI	TLI	RMSEA	SRMR
Model 1 Assertions-Overall (one-factor model)	**1**	0.977	0.976	**0.094**	**0.130**
	2	0.974	0.973	**0.082**	**0.118**
Model 2 Assertions-Disciplines - Overall (second-order model)	**1**	0.986	0.986	**0.072**	**0.106**
	2	0.985	0.984	**0.062**	**0.094**
Model 3 Assertions-Clusters-Overall (second-order model)	**1**	0.992	0.991	0.057	**0.083**
	2	0.991	0.991	0.048	0.072
Model 4 Assertions-Clusters-Disciplines-Overall (third-order model)	**1**	0.992	0.991	0.057	**0.083**
	2	0.991	0.991	0.048	0.072

Note. Numbers in bold do not meet the criteria for goodness of fit.

Table 17 shows the estimated correlations among disciplines for Model 4 (third-order model). The correlations are all very high, ranging between 0.913 and 1. The high correlations between the disciplines in Model 4 indicate that, after accounting for the cluster effects, the disciplines do not add much to the model. This may explain why Model 4 did not show an improvement in fit compared to Model 3. Overall, the findings support the IRT model used for calibration.

Table 17. Model-Implied Correlations per Form for the Disciplines in Model 4

Grade	Form	Discipline	Earth and Space Sciences	Life Sciences
6	**1**	Physical Sciences	0.999	0.941
		Earth and Space Sciences	–	0.940
	2	Physical Sciences	1.000	0.964
		Earth and Space Sciences	–	0.964
	3	Physical Sciences	0.975	0.923
		Earth and Space Sciences	–	0.947
7	**1**	Physical Sciences	0.983	0.947
		Earth and Space Sciences	–	0.937
	2	Physical Sciences	0.978	0.972
		Earth and Space Sciences	–	0.951
	3	Physical Sciences	0.955	0.936
		Earth and Space Sciences	–	0.966
	4	Physical Sciences	0.938	0.913
		Earth and Space Sciences	–	0.973
	5	Physical Sciences	0.931	0.944
		Earth and Space Sciences	–	0.965
	6	Physical Sciences	0.941	0.928

Grade	Form	Discipline	Earth and Space Sciences	Life Sciences
		Earth and Space Sciences	–	0.967
8	1	Physical Sciences	0.971	0.971
		Earth and Space Sciences	–	0.970
	2	Physical Sciences	0.956	0.958
		Earth and Space Sciences	–	0.935
	3	Physical Sciences	0.966	0.978
		Earth and Space Sciences	–	0.988

5.4.2 Conclusion

The models with no cluster effects provided the highest degrees of misfit across forms and grades (Models 1 and 2), indicating that the cluster effects must be taken into account as additional latent variables. On the other hand, once the cluster effects were accounted for, a single science dimension was sufficient (Model 3): including additional dimensions for the science disciplines (Life Science, Physical Science, Earth and Space Sciences) did not improve model fit and the correlations among those three dimensions were very high (Model 4). Model 3, with a single overall dimension for science and additional latent variables to account for the effect of item clusters, provided the best balance between model fit and parsimony.

Overall, the findings support the use of the Rasch testlet model as the IRT calibration model and the reporting of an overall score directly computed from all the items a student took. Because there are enough items within each discipline in the test blueprint, discipline subscores can be reported at the individual level although they may not provide much unique information from the total score for most students. However, many stakeholders often desire information about student performance in addition to a single overall score. Note that it is not uncommon to provide subscores at the individual level even when the assessment is essentially unidimensional in a psychometric sense. For example, evidence derived from the dimensionality analyses for the Smarter Balanced Assessment suggests that "no consistent and pervasive multidimensionality was demonstrated" (Smarter Balanced Assessment Consortium, 2016, p. 182), yet individual claim scores are routinely reported in addition to overall ELA and mathematics scores.

6. Fairness in Content

The principles of universal design (UD) provide guidelines for test design that minimize the impact of construct-irrelevant factors in assessing student performance. UD removes barriers in order to provide access for the widest possible range of students. The following seven principles of UD are applied in the process of test development (Thompson, Johnstone, & Thurlow, 2002):

1. Inclusive assessment population
2. Precisely defined constructs

3. Accessible, non-biased items
4. Amenable to accommodations
5. Simple, clear, and intuitive instructions and procedures
6. Maximum readability and comprehensibility
7. Maximum legibility

Test development specialists have received extensive training on UD principles and applied them in the development of all test materials. During the review process, adherence to UD principles is verified by Montana educators and stakeholders. More details on how to reduce construct-irrelevant variance through universal design and on training on the principles of universal design are described in Section 2, Item Development Process That Supports Validity of Claims, as well as Appendix 2-C, Style Guide for Science Items, of Volume 2 of this technical report.

6.1 Cognitive Laboratory Studies

In 2017, when the development of item clusters for the states that are part of the Memorandum of Understanding (MOU) began, cognitive lab studies were carried out to evaluate and refine the process of developing item clusters aligned to the Next Generation Science Standards (NGSS). Results of the cognitive lab studies confirmed the feasibility of the approach. Item clusters were completed within 12 minutes on average, and students reported being familiar with the format conventions and online tools used in the item clusters. They appeared to easily navigate the item clusters' interactive features and response formats. In general, students who received credit on a given item displayed a reasoning process that aligned with the skills that the item was intended to measure.

A second set of cognitive lab studies were carried out in 2018 and 2019 to determine if students using braille can understand the task demands of selected accommodated three-dimensional science standards–aligned item clusters and can navigate the interactive features of these clusters in a manner that allows them to fully display their knowledge and skills relative to the constructs of interest. In general, both students who relied entirely on braille and/or the Job Access with Speech (JAWS) screen-reading software and those who had some vision and were able to read the screen with magnification were able to find the information they needed to respond to the questions, navigate the various response formats, and finish within a reasonable amount of time. The clusters were clearly different from (and more complex than) other tests with which the students were familiar, however; the study therefore recommended that students be given adequate time to practice with at least one sample cluster before taking the summative test. The study also resulted in tool-specific recommendations to increase accessibility for visually impaired students. The reports of both sets of cognitive lab studies are presented in Appendix 4-D, Science Clusters Cognitive Lab Report, and Appendix 4-E, Braille Cognitive Lab Report.

6.2 Statistical Fairness in Item Statistics

Differential item functioning (DIF) analyses were conducted with other states that field-tested the items for the initial item bank. A thorough content review was performed in those states. The details surrounding this review of items for bias is further described in Section 4.4 of Volume 1, Annual Technical Report, along with the DIF analysis process for the Montana Science Assessment.

7. SUMMARY

This volume is intended to provide a collection of reliability and validity evidence that supports appropriate inferences from the observed test scores. The overall results can be summarized as follows:

- ***Reliability.*** Various measures of reliability are provided at the aggregate and subgroup levels, showing that the reliability of all tests is in line with acceptable industry standards.
- ***Content Validity.*** Evidence is provided to support the assertion that content coverage on each test was consistent with the test specifications of the blueprint across testing modes.
- ***Internal Structural Validity.*** Evidence is provided to support the selection of the measurement model, the tenability of model assumptions, and the reporting of an overall score and subscores at the reporting category levels.
- ***Relationship of Test Scores to External Variables.*** Evidence of convergent and discriminant validity is provided to support the relationship between the test and other measures intended to assess similar constructs, as well as between the test and other measures intended to assess different constructs.
- ***Test Fairness.*** Items are developed in accordance with the principles of universal design, which removes barriers in order to provide access for the widest possible range of students. Evidence of test fairness is provided statistically using differential item functioning analysis in tandem with content reviews by specialists.

8. REFERENCES

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

Bradlow, E. T., Wainer, H., & Wang, X. (1999). A Bayesian random effects model for testlets. *Psychometrika*, *64*, 153–168.

Brown, T. A. (2015). *Confirmatory factor analysis for applied research* (2nd ed.). New York: The Guilford Press.

Crocker, L., & Algina, J. (1986). *Introduction to classical and modern test theory*. New York: Holt, Rinehart, and Winston.

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, *6*(1), 1–55.

Li, Y., Bolt, D. M., & Fu, J. (2006). A comparison of alternative models for testlets. *Applied Psychological Measurement*, *30*, 3–21.

Messick, S. (1989). Validity. In R. L. Linn (Ed.), *Educational Measurement* (3rd ed., pp. 13–103). New York: Macmillan.

Rijmen, F. (2009). *Three multidimensional models for testlet-based tests: Formal relations and an empirical comparison*. (ETS Research Rep. No. RR-09-37). Princeton, NJ: ETS.

Rijmen, F., Jiang, T., & Turhan, A. (2018, April). An item response theory model for new science assessments. Paper presented at the annual meeting of the National Council on Measurement in Education, New York, NY.

Smarter Balanced Assessment Consortium. (2016). *2013–2014 Technical Report.* Retrieved from https://portal.smarterbalanced.org/library/en/2013-14-technical-report.pdf.

Thompson, S. J., Johnstone, C. J., & Thurlow, M. L. (2002). *Universal design applied to large scale assessments*. (Synthesis Report 44). Minneapolis: University of Minnesota, National Center on Educational Outcomes. Retrieved from http://education.umn.edu/NCEO/OnlinePubs/Synthesis44.html.

Wang, W. C., & Wilson, M. (2005). The Rasch testlet model. *Applied Psychological Measurement*, *29*(2), 126–149.

Yung, Y. F., Thissen, D., & McLeod, L. D. (1999). On the relationship between the higher-order factor model and the hierarchical factor model. *Psychometrika*, *64*, 113–128.

Appendix 4-A

Student Demographics and Reliability Coefficients

Student Demographics and Reliability Coefficients

Table A-1. Marginal Reliability Coefficients by Demographic Subgroups

Group	Grade 5	Grade 8
All Students	0.88	0.89
Female	0.88	0.88
Male	0.89	0.90
African American	0.88	0.87
American Indian/Native Alaskan	0.87	0.90
Asian	0.87	0.89
Hawaiian/Pacific Islander	0.88	0.88
Hispanic	0.87	0.88
Multi-Racial	0.84	0.85
White	0.88	0.89
Limited English Proficiency	0.84	0.86
Special Education	0.86	0.86
Economically Disadvantaged	0.87	0.88

Table A-2. Scale Score Summary by Reporting Category, Science Grade 5

Reporting Category	Mean	SD	Min	Max	Reliability	SEM
Physical Sciences	498.33	30.61	400.12	599.83	0.70	16.58
Earth and Space Sciences	497.43	33.98	400.12	599.83	0.75	16.67
Life Sciences	499.62	31.75	400.12	599.83	0.70	17.28

Table A-3. Scale Score Summary by Reporting Category, Science Grade 8

Reporting Category	Mean	SD	Min	Max	Reliability	SEM
Physical Sciences	801.60	30.25	700.10	899.82	0.73	15.55
Earth and Space Sciences	799.15	33.26	700.10	899.82	0.73	16.99
Life Sciences	799.40	31.31	700.10	899.82	0.73	16.19

Appendix 4-B

Conditional Standard Error of Measurement

Conditional Standard Error of Measurement

Table B-1. Spring 2023 CSEM at Each Scale Score, Science Grade 5

Science Grade 5

Scale Score	Performance Level	CSEM
400	1	22.00
402	1	17.40
411	1	14.94
413	1	16.61
414	1	14.46
415	1	14.97
416	1	15.07
419	1	12.82
420	1	13.53
422	1	13.71
423	1	13.47
424	1	12.74
425	1	13.55
426	1	13.37
427	1	12.09
428	1	12.68
429	1	12.64
430	1	11.89
431	1	12.02
432	1	12.10
433	1	12.06
434	1	11.46
435	1	11.84
436	1	12.10
437	1	11.66
438	1	11.85
439	1	11.32
440	1	11.31
441	1	11.01
442	1	11.30
443	1	10.89
444	1	11.00
445	1	10.86

Science Grade 5

Scale Score	Performance Level	CSEM
446	1	10.92
447	1	10.67
448	1	10.63
449	1	10.42
450	1	10.52
451	1	10.22
452	1	10.14
453	1	10.29
454	1	10.13
455	1	10.07
456	1	10.17
457	1	9.97
458	1	9.79
459	1	9.89
460	1	9.78
461	1	9.80
462	1	9.78
463	1	9.58
464	1	9.56
465	1	9.65
466	1	9.58
467	1	9.49
468	1	9.48
469	1	9.44
470	1	9.44
471	1	9.45
472	1	9.33
473	1	9.34
474	1	9.19
475	1	9.32
476	1	9.33
477	2	9.23
478	2	9.25
479	2	9.22
480	2	9.27
481	2	9.20

Science Grade 5

Scale Score	Performance Level	CSEM
482	2	9.20
483	2	9.21
484	2	9.18
485	2	9.16
486	2	9.19
487	2	9.22
488	2	9.11
489	2	9.20
490	2	9.14
491	2	9.17
492	2	9.22
493	2	9.05
494	2	9.20
495	2	9.11
496	2	9.04
497	2	9.06
498	2	9.05
499	2	9.14
500	2	9.03
501	2	9.06
502	2	9.07
503	2	9.10
504	2	9.08
505	2	9.10
506	3	9.11
507	3	9.04
508	3	9.10
509	3	9.10
510	3	9.08
511	3	9.05
512	3	9.06
513	3	9.07
514	3	9.04
515	3	9.15
516	3	9.04
517	3	9.05

Science Grade 5

Scale Score	Performance Level	CSEM
518	3	9.01
519	3	9.20
520	3	9.02
521	3	9.14
522	3	9.11
523	3	9.07
524	3	9.10
525	3	9.13
526	3	9.14
527	3	9.19
528	3	9.16
529	3	9.21
530	3	9.22
531	4	9.21
532	4	9.27
533	4	9.33
534	4	9.29
535	4	9.41
536	4	9.36
537	4	9.34
538	4	9.35
539	4	9.44
540	4	9.45
541	4	9.53
542	4	9.54
543	4	9.67
544	4	9.59
545	4	9.66
546	4	9.78
547	4	9.97
548	4	10.02
549	4	9.86
550	4	9.94
551	4	10.08
552	4	10.14
553	4	10.07

Science Grade 5

Scale Score	Performance Level	CSEM
554	4	10.08
555	4	10.42
556	4	10.36
557	4	10.20
558	4	10.40
559	4	10.71
560	4	10.56
561	4	10.86
562	4	11.15
563	4	10.97
564	4	11.09
565	4	11.28
566	4	10.94
567	4	10.97
568	4	11.61
569	4	11.44
570	4	12.10
571	4	11.28
572	4	11.27
573	4	11.62
574	4	11.51
575	4	11.70
577	4	12.52
578	4	11.73
579	4	12.30
580	4	12.64
581	4	12.01
584	4	13.48
585	4	13.94
587	4	14.10
588	4	13.25
589	4	12.43
599	4	14.25
600	4	16.65

Table B-2. Spring 2023 CSEM at Each Scale Score, Science Grade 8

Science Grade 8

Scale Score	Performance Level	CSEM
700	1	22.89
702	1	19.15
709	1	16.07
713	1	14.96
716	1	14.94
718	1	13.44
720	1	13.54
721	1	13.89
722	1	13.61
723	1	13.09
724	1	12.77
725	1	13.40
726	1	13.46
728	1	12.23
729	1	12.76
730	1	12.44
731	1	12.29
732	1	11.95
733	1	11.95
734	1	12.00
735	1	11.79
736	1	11.73
737	1	11.53
738	1	11.49
739	1	11.39
740	1	11.31
741	1	11.18
742	1	11.02
743	1	11.07
744	1	10.96
745	1	10.79
746	1	10.78
747	1	10.72
748	1	10.73
749	1	10.60

Science Grade 8

Scale Score	Performance Level	CSEM
750	1	10.60
751	1	10.46
752	1	10.42
753	1	10.38
754	1	10.22
755	1	10.16
756	1	10.04
757	1	10.09
758	1	10.02
759	1	10.02
760	1	9.95
761	1	9.86
762	1	9.82
763	1	9.75
764	1	9.72
765	1	9.67
766	1	9.63
767	1	9.61
768	1	9.60
769	1	9.59
770	1	9.59
771	1	9.42
772	1	9.43
773	1	9.43
774	1	9.44
775	1	9.31
776	1	9.37
777	1	9.40
778	1	9.24
779	1	9.37
780	1	9.25
781	2	9.23
782	2	9.31
783	2	9.24
784	2	9.17
785	2	9.08

Science Grade 8

Scale Score	Performance Level	CSEM
786	2	9.09
787	2	9.12
788	2	9.08
789	2	9.01
790	2	8.98
791	2	8.87
792	2	8.88
793	2	8.90
794	2	8.85
795	2	8.78
796	2	8.76
797	2	8.75
798	2	8.70
799	2	8.73
800	2	8.70
801	2	8.70
802	2	8.62
803	2	8.61
804	2	8.57
805	2	8.57
806	2	8.63
807	2	8.59
808	3	8.57
809	3	8.57
810	3	8.64
811	3	8.60
812	3	8.61
813	3	8.58
814	3	8.61
815	3	8.58
816	3	8.65
817	3	8.61
818	3	8.63
819	3	8.71
820	3	8.71
821	3	8.74

Science Grade 8

Scale Score	Performance Level	CSEM
822	3	8.79
823	3	8.83
824	3	8.83
825	3	8.84
826	4	8.81
827	4	8.91
828	4	8.90
829	4	8.90
830	4	8.88
831	4	8.86
832	4	8.88
833	4	9.02
834	4	9.06
835	4	8.96
836	4	9.10
837	4	9.03
838	4	9.06
839	4	9.07
840	4	9.09
841	4	9.15
842	4	9.03
843	4	9.17
844	4	9.07
845	4	9.34
846	4	9.29
847	4	9.31
848	4	9.24
849	4	9.46
850	4	9.26
851	4	9.43
852	4	9.48
853	4	9.43
854	4	9.31
855	4	9.54
856	4	9.48
857	4	9.50

Science Grade 8

Scale Score	Performance Level	CSEM
858	4	9.55
859	4	9.54
860	4	9.79
861	4	9.55
862	4	9.71
863	4	9.79
864	4	9.67
865	4	9.68
866	4	9.84
867	4	9.66
868	4	9.93
869	4	9.85
870	4	9.56
871	4	9.93
872	4	9.98
873	4	9.98
874	4	9.98
875	4	10.05
876	4	10.17
877	4	10.56
878	4	10.02
879	4	10.48
881	4	11.10
882	4	11.03
884	4	10.65
885	4	10.42
886	4	10.01
887	4	10.90
888	4	11.25
890	4	10.96
891	4	10.86
892	4	11.42
893	4	11.15
895	4	11.93
896	4	10.98
900	4	12.42

Appendix 4-C

Classification Accuracy and Consistency Indices by Subgroups

Classification Accuracy and Consistency Indices by Subgroups

Table C-1. Spring 2023 Classification Accuracy by Demographic Subgroup

Group	N	Overall (%)	By Cut (%)			By Level (%)			
			Level 2 Cut	*Level 3 Cut*	*Level 4 Cut*	*Level 1*	*Level 2*	*Level 3*	*Level 4*
Grade 5									
All Students	11,033	76.98	92.55	90.18	94.22	85.81	75.65	70.62	79.98
Female	5,391	76.86	92.49	89.84	94.51	85.80	75.73	70.69	78.95
Male	5,642	77.10	92.61	90.50	93.95	85.83	75.56	70.54	80.78
African American	88	78.92	90.19	93.69	95.02	87.10	74.53	76.09	81.98
American Indian/Alaskan Native	1,079	81.97	88.96	94.26	98.74	89.00	75.40	70.94	77.27
Asian	80	75.36	94.33	90.02	90.97	80.52	75.23	72.45	77.76
Hispanic	606	77.10	91.33	90.22	95.52	87.38	75.58	69.75	76.72
Multi-Racial	533	77.08	91.87	89.91	95.28	87.41	74.40	70.22	80.29
Pacific Islander	16	77.68	95.04	90.72	91.88	93.55	80.69	74.05	53.92
White	8,631	76.34	93.13	89.65	93.53	84.36	75.77	70.61	80.17
Limited English Proficiency (LEP)	372	81.67	89.24	93.88	98.54	89.15	75.46	69.25	73.56
Non-LEP	10,661	76.82	92.67	90.05	94.07	85.52	75.65	70.64	80.01
Special Education (SPED)	1,442	82.59	89.20	95.05	98.33	89.23	75.88	70.46	80.46
Non-SPED	9,591	76.14	93.05	89.45	93.61	84.21	75.61	70.63	79.96
Economically Disadvantaged	4,831	78.57	90.80	91.37	96.38	87.22	75.88	70.13	79.27
Non-Economically Disadvantaged	6,202	75.74	93.92	89.25	92.55	83.17	75.44	70.86	80.20

Group	N	Overall (%)	By Cut (%)			By Level (%)			
			Level 2 Cut	*Level 3 Cut*	*Level 4 Cut*	*Level 1*	*Level 2*	*Level 3*	*Level 4*
Grade 8									
All Students	11,112	76.09	93.44	88.95	93.53	88.47	74.57	61.88	83.12
Female	5,359	75.07	93.33	88.21	93.36	87.76	74.60	61.88	81.16
Male	5,753	77.04	93.55	89.65	93.69	89.03	74.54	61.89	84.58
African American	56	76.81	90.46	88.21	98.04	83.62	75.61	62.40	93.77
American Indian/Alaskan Native	1,076	80.77	89.71	93.11	97.87	89.85	73.96	60.87	78.62
Asian	77	74.40	92.59	88.64	93.01	86.23	73.22	63.85	80.23
Hispanic	656	77.65	91.95	90.24	95.33	89.23	74.29	60.36	84.83
Multi-Racial	529	76.92	93.36	88.96	94.46	90.44	76.01	61.25	83.47
Pacific Islander	24	78.56	89.34	92.38	96.71	87.69	76.09	58.66	78.51
White	8,694	75.34	94.06	88.34	92.77	87.84	74.57	62.03	83.13
Limited English Proficiency (LEP)	451	83.79	90.91	94.55	98.27	91.84	74.44	58.57	86.53
Non-LEP	10,661	75.76	93.55	88.72	93.33	88.07	74.57	61.93	83.09
Special Education (SPED)	1,284	82.91	89.84	94.74	98.26	91.15	73.97	61.02	82.26
Non-SPED	9,828	75.20	93.91	88.20	92.91	87.40	74.64	61.92	83.14
Economically Disadvantaged	4,041	78.00	91.24	90.72	95.91	89.11	74.44	61.23	80.65
Non-Economically Disadvantaged	7,071	74.99	94.70	87.94	92.17	87.61	74.65	62.10	83.70

Table C-2. Spring 2023 Classification Consistency by Demographic Subgroup

Group	N	Overall (%)	By Cut (%)			By Level (%)			
			Level 2 Cut	*Level 3 Cut*	*Level 4 Cut*	*Level 1*	*Level 2*	*Level 3*	*Level 4*
Grade 5									
All Students	11,033	68.00	89.43	86.28	91.85	77.10	66.82	61.08	69.65
Female	5,391	67.82	89.30	85.81	92.25	76.90	67.14	61.06	67.85
Male	5,642	68.18	89.55	86.72	91.47	77.29	66.50	61.10	71.05
African American	88	69.93	86.39	90.09	93.05	76.70	68.45	63.63	69.85
American Indian/Alaskan Native	1,079	74.47	84.28	91.83	98.09	84.75	66.01	58.30	51.66
Asian	80	65.44	91.76	85.92	87.24	72.21	64.13	64.17	64.66
Hispanic	606	68.39	87.71	86.52	93.72	77.88	66.98	61.53	63.97
Multi-Racial	533	68.49	88.66	86.16	93.24	78.28	67.28	59.76	69.20
Pacific Islander	16	68.37	91.70	86.32	89.74	80.51	70.94	64.39	43.65
White	8,631	67.14	90.24	85.54	90.89	74.36	66.87	61.20	70.26
Limited English Proficiency (LEP)	372	74.36	84.88	91.35	97.87	85.23	65.30	58.31	54.56
Non-LEP	10,661	67.78	89.59	86.10	91.64	76.44	66.87	61.12	69.74
SPED	1,442	75.33	84.65	92.84	97.60	85.40	66.25	56.61	67.63
Non-SPED	9,591	66.90	90.15	85.29	90.99	73.57	66.90	61.32	69.72
Economically Disadvantaged	4,831	70.03	86.87	87.93	94.85	80.36	67.25	59.86	66.53
Non-Economically Disadvantaged	6,202	66.42	91.42	84.99	89.52	71.43	66.46	61.69	70.69
Grade 8									
All Students	11,112	67.39	90.66	84.69	90.83	80.59	65.75	51.15	73.87

Group	N	Overall (%)	By Cut (%)			By Level (%)			
			Level 2 Cut	*Level 3 Cut*	*Level 4 Cut*	*Level 1*	*Level 2*	*Level 3*	*Level 4*
Female	5,359	66.01	90.47	83.65	90.61	78.75	66.08	51.20	70.41
Male	5,753	68.67	90.83	85.66	91.03	82.09	65.40	51.10	76.54
African American	56	68.31	86.64	84.14	96.54	81.42	64.44	51.89	66.04
American Indian/Alaskan Native	1,076	73.55	85.61	90.38	96.92	85.88	65.38	47.55	58.06
Asian	77	65.26	89.82	84.08	89.97	77.36	61.19	55.39	72.13
Hispanic	656	69.45	88.69	86.45	93.32	83.07	66.31	48.43	72.66
Multi-Racial	529	68.18	90.21	84.68	92.12	80.88	67.89	50.65	72.65
Pacific Islander	24	70.25	84.05	89.94	95.56	79.51	66.20	47.26	78.69
White	8,694	66.43	91.51	83.85	89.77	78.60	65.65	51.49	74.35
Limited English Proficiency (LEP)	451	77.61	87.01	92.51	97.58	88.96	65.04	45.82	68.48
Non-LEP	10,661	66.96	90.81	84.36	90.55	79.67	65.77	51.23	73.91
SPED	1,284	76.23	85.64	92.58	97.49	87.29	65.24	46.42	68.97
Non-SPED	9,828	66.24	91.31	83.66	89.96	78.09	65.80	51.36	74.00
Economically Disadvantaged	4,041	69.84	87.59	87.14	94.19	83.10	66.34	49.01	69.30
Non-Economically Disadvantaged	7,071	65.99	92.40	83.29	88.91	77.45	65.39	51.93	74.98

Appendix 4-D

Science Clusters Cognitive Lab Report

Science Cluster Cognitive Interviews

Fran Stancavage

Susan Cole

March 2018

TABLE OF CONTENTS

LIST OF TABLES

List of Figures

1. INTRODUCTION

American Institutes for Research (AIR) and a group of states are developing methods to measure student learning of Next Generation Science Standards (NGSS) and other standards derived from the K–12 science framework. Educators involved in the development of the framework and the standards encourage measuring learning using integrated tasks that require a student's sustained concentration on a realistic science or engineering task. This set of cognitive interviews was undertaken early in the development process to test and refine our approach to developing item clusters to measure NGSS and related performance expectations (PEs).

The approach taken for each cluster was to identify a *phenomenon* to be explained, modeled, described, or analyzed (as appropriate for the performance expectation) and have a sequence of interrelated, often interdependent items (some containing multiple interactions) that build to support the completion of a task.

This set of cognitive interviews was designed to provide data on newly developed item clusters aligned with the NGSS. We evaluated 12 clusters, four designed for elementary school, four designed for middle school, and four designed for high school. Each cluster contained one to five items, many with separately scored sub-items. Per the request of the item development team, the labs focused on the following questions:

- How long did students take to respond to each cluster?
- How well did students score on each item and on each cluster overall?
- What aspects of the items were confusing to students?
- What reasoning skills did students display as they worked their way through each item?

A limitation of the cognitive lab analysis was that many of the students had limited exposure to content covered in the clusters, particularly the clusters on German Pyramid Candle (elementary school), Morning Fog (middle school), Texas Weather (middle school), Saving the Tuna (high school), and Tomcods (high school). To partially offset this lack of formal instruction, students were provided with a one- or two-page hard-copy lesson on the relevant science content for each cluster. Some of the later cognitive interviews were conducted in schools in which the teachers had received substantial training in teaching the new standards.

The remainder of this report includes an overview of methods, a description of the study sample, a discussion of the findings for each of the 12 clusters, and a final section on the students' overall perceptions of the science clusters.

2. METHODS

2.1 STUDY DESIGN

Between January and May 2017, cognitive interviews were conducted with 18 elementary school students, 12 middle school students, and 15 high school students. The interviews lasted one and one-half hours, and each student was presented with all four clusters for their grade level. The order of the clusters was rotated so that the risk of student fatigue or missing responses was distributed across the clusters.

Students were encouraged to think out loud while they were responding to the items (concurrent think-aloud), and interviewers were instructed to use follow-up probes to clarify and expand on what each student said (or what each student was observed to do). To preclude the possibility that students' responses to later items would be influenced by probing on earlier items, probes were only administered after students had completed all the items in a cluster.

At the start of the interview, the interviewer trained the student on the concurrent think-aloud technique. The interviewer first modeled the technique and then had the student practice on one or, if necessary, two items. Lower grade multiple-choice mathematics items were used for the modeling and practice.

After the think-aloud training, students were provided with a hard-copy lesson on the relevant science content, as described previously. The item development team developed the lessons, and the interviewer collected the hard copy before the student stared the cluster.

At the end of the cognitive interview, each student was asked three general questions: (1) whether the student had studied any of the cluster topics in school, (2) whether the student had taken tests that look similar and/or used similar tools, and (3) how hard the student thought this test was.

2.2 TRAINING AND PILOT TESTING

Five interviewers (and one backup interviewer) were trained for the project. Since all the interviewers were experienced in the cognitive interview technique, the training primarily focused on reviewing the content of the clusters and familiarizing the interviewers with the test platform and the specifics of the interview protocols. Project leads provided a separate two-hour training for the protocol at each grade level.

Additionally, at each grade level, an experienced team member conducted a pilot interview to fine tune the protocol and, especially, to determine the number of clusters that could be covered in one interview and hence the number of students that would be required to adequately test the clusters. The pilot administrations confirmed that, at each grade level, all four clusters could be covered in a single one and one-half hour interview. Thus, for each cluster, we ultimately had data on 12 to 18 students.

2.3 STUDY SAMPLE

Students were primarily drawn from the San Francisco Bay area. Utah also contributed students for the elementary school sample, and Connecticut contributed students for the high school sample.

The Utah students were particularly valuable to the study because they were in schools where teachers were receiving Next Generation Science Standards (NGSS) training from an NGSS author.

To recruit students in the San Francisco Bay area, the project manager and a designated scheduler at the American Institutes for Research (AIR) worked with a recruitment firm. This firm used a household-based approach to recruitment and employed an AIR-developed recruitment screener. Having recognized that exposure to inquiry-based science would be limited, we targeted higher achieving students with the expectation that they would be the most likely to have received this instruction and have benefited from it. We tried to recruit students whose parents reported the students' grades as being mostly As and/or Bs in science. We balanced the sample on gender and ethnicity (white/non-white).

In Utah and Connecticut, the AIR program manager worked directly with designated school districts to recruit students near Salt Lake City and Hartford, respectively. The cognitive interviews were conducted at the AIR offices in San Mateo, California, and on-site at the schools in Utah and Connecticut. The characteristics of the sample are summarized in Table 1 and shown by student in the Appendix.

Table 1. Characteristics of Sample, by Grade Level

Characteristic	Elementary School (*n* = 18)	Middle School (*n* = 12)	High School (*n* = 15)
Location			
California	12	12	12
Connecticut	N/A	N/A	3
Utah	6	N/A	N/A
Grade Level			
Grade 5	15	N/A	N/A
Grade 6	3[1]	N/A	N/A
Grade 8	N/A	7	N/A
Grade 9	N/A	5	N/A
Grade 10	N/A	N/A	1[2]
Grade 11	N/A	N/A	13
Grade 12	N/A	N/A	1[2]
Gender			
Male	13	6	5
Female	5	6	10
Parent or Teacher Reported Ethnicity			
African American	1	2	1
Asian	2	3	1
Hispanic	1	1	5
White	13	6	6

Characteristic	Elementary School (*n* = 18)	Middle School (*n* = 12)	High School (*n* = 15)
Other	1	0	1
Prefer not to answer	0	0	1
Parent-Reported Achievement in Science[3]			
Mostly As	7	11	7
Mostly Bs	5	1	5

[1] Utah students
[2] Connecticut students
[3] Data for California subjects only

3. FINDINGS

We begin this section with a summary of findings that includes key take-aways from the cognitive interviews and basic performance statistics for each of the 12 clusters.

The summary is followed by a detailed discussion of cognitive interview findings for each of the 12 clusters. Each cluster-level discussion starts with a summary of student performance, a list of task demands, and an image of the cluster stimulus. These are followed by an item-by-item discussion that, for each item, displays the item text, summarizes score patterns, and addresses students' comprehension and reasoning.

The discussion of findings ends with a summary of students' general perceptions of the science clusters, as expressed at the end of the cognitive interviews.

3.1 SUMMARY OF FINDINGS

3.1.1 Key Take-Aways

Feasibility of Cluster Approach

Results from the cognitive interviews suggest that it is feasible to incorporate item clusters into standardized science tests. On average, the clusters took 12 minutes to complete, and students reported being familiar with the format conventions and tools used in the clusters and appeared to easily navigate the clusters' interactive features and response formats.

- When questioned at the end of the cognitive interviews, nearly all students at each grade level reported that they had taken online tests that used similar page layouts, multimedia, and tools (e.g., page layouts with stimulus on the left and items on the right; embedded video; scroll bars; Back, Next, and Zoom in/Zoom out buttons; drop-down menus; and connect line and Add Arrow tools).
- Further, interviewers noted that students at all grade levels appeared comfortable navigating the clusters and, generally speaking, understood how to interact with the simulations and the response formats. When students experienced confusion, it was due to idiosyncratic problems with specific simulations or test items.

Relationship to Content Knowledge

Across grade levels, most students who participated in the cognitive interviews found the greatest challenge to be their lack of relevant content knowledge or experience applying science and engineering practices. This is not unexpected given that the clusters were built to measure NGSS constructs, and most of the students in the sample had not been exposed to NGSS-based instruction.

- Utah students, who were specifically included in the elementary school sample because they came from schools in which teachers were receiving NGSS training from an NGSS author, did better on all clusters. Details are given in the next subsection, where we summarize student performance by cluster.

Many students commented on their lack of relevant content knowledge during the think-alouds, and, when questioned at the end of the interview, students reported that they lacked prior instruction in most of the topics covered by the clusters. If they had studied those topics, they said that it was at less depth than required to be successful. For example, one high school student said, in reference to the Blood Sugar Regulation cluster, that she had reviewed molecule concentrations but never discussed how they are impacted by meals, "not that in-depth, more gone over these and what they do for the body."

- By contrast, one of the Utah students said he had studied all four elementary school topics. "At the beginning of the year we studied the heat one and how we can help make a motor turn something on, like a light bulb. I thought of that. Maybe it was just backwards, the light was helping the fan to spin. The light was turning or making it spin by the energy it was producing. I remember last year, in 4th grade, we studied the Grand Canyon and the animals, and we did a little bit this year, and the animals that were living in the walls like trilobite and some others like starfish. We saw this video of this hole that was in Arizona, and there were tons of fossils in it. I think we studied a little bit on the terrarium one . . . We studied a little bit about [the desert plants]. About how each plant could survive."

Measuring Intended Constructs

In general, students who received credit on a given item (and some who did not) displayed a reasoning process that aligned with the skills that the item was intended to measure.

- This held true even for standard multiple-choice or multi-select items. For example, thinking aloud as he responded to this question in the Redwall Limestone cluster,

> **Part A**
>
> Within the Grand Canyon, a rock layer contains fossils of octopi (plural of "octopus"), brachiopods, and corals. What can you conclude about the environment of the Grand Canyon region from the fossil evidence?
>
> Ⓐ The Grand Canyon region was always desert.
>
> Ⓑ The Grand Canyon region was once underwater.
>
> Ⓒ The Grand Canyon region experienced a lot of rain.
>
> Ⓓ The fossils do not provide any information about the environment.

one elementary school student first read option A, *[t]he Grand Canyon region was always desert*, out loud. Then he said he wanted to check the next option and read *[t]he Grand Canyon region was once underwater*. The student said that option B could be the answer, "but the first option [A] is not because it said in the question [the fossils] were sea animals." The student then read option C, *[t]he Grand Canyon region experienced a lot of rain*, and option D, *[t]he fossils do not provide any information about the environment*. He said that the answer couldn't be option D because "[the question] doesn't have anything to do with the animals that are living today." He said it probably wasn't option C because "even if it rained, [but] it wasn't an ocean, then the coral couldn't live there." The student concluded that the correct answer had to be B.

- In another example, an elementary school student explained her response to Part B of this two-part item from the Desert Plants cluster

The following question has two parts. First, answer part A. Then, answer part B.

Use the data from the experiment to compare the survival of the three types of plants in the desert.

Part A

Record the data from the experiment by adding numbers to the table.

	Mesquite Trees	Cactus Plants	Bird's Nest Ferns
Number of plants at start of experiment			
Number of plants at end of experiment			

Part B

Select the **two** statements that are supported by the data in the table you created.

☐ All types of plants can survive in all environments.

☐ No types of plants can survive in a dry desert environment.

☐ All types of plants can survive in the dry desert environment.

☐ Some types of plants cannot survive in the dry desert environment.

☐ Some types of plants survive better than others in the dry desert environment.

by saying that she chose the second-to-last option (*[s]ome types of plants cannot survive in the dry desert environment*) because "at the start of the experiment, there was a total of 5 bird's nest ferns, and then they all died, and also because one of the mesquite trees – they died – but I mean, most of them still remained." And she chose the last option (*[s]ome types of plants survive better than others in the dry desert environment*) because "out of all 3 of the plants, the cactus all lived instead of dying." She shared that she did not choose the first option (*[a]ll types of plants can survive in all environments*) because "As you can see, some of them died – like the bird's nest ferns and the mesquite trees." She shared that she did not choose the second option (*[n]o types of plants can survive in a dry desert environment*) "because the cactus – they still lived." She shared that she did not choose the third option (*[a]ll types of plants can survive in the dry desert environment*) "because the bird's nest ferns died."

There were exceptions where students gained or lost credit for non-construct relevant reasons, but these were related to specific item flaws that could be fixed before the items were used operationally.

General Recommendations for Improvements

While the validity of the general approach was supported by the cognitive lab findings, there were flaws in specific types of items that can and should be remediated before using the items operationally:

- Students needed more cueing on multi-select items such as the following:

Part B

From the list of additional experiments, select the evidence that would support your answer in part A.

- ☐ Scientists grow a sample of wild-type *Mycobacterium tuberculosis* in the lab. Over time, some of the bacteria show resistance to rifampin.
- ☐ Scientists plate a colony of wild-type *Mycobacterium tuberculosis* and a colony of *Escherichia coli* in one petri dish. Some of the new colonies show resistance to rifampin.
- ☐ Scientists plate a colony of wild-type *Mycobacterium tuberculosis* and a colony of mutant *Mycobacterium tuberculosis* in one petri dish. Some of the new colonies show resistance to rifampin.
- ☐ Scientists create additional *Mycobacterium tuberculosis* mutants by creating substitution mutations in the DNA that codes for amino acids 36-67. Many of the mutants are resistant to rifampin.

Earning a score point for this item required correctly selecting both the first and the last options, but most students stopped after choosing one response. This type of error could be minimized by adding “mark all that apply” to the item stem.

- Students interactions with simulations should be checked to make sure that the simulations are functioning as intended. For example, a flaw in the simulation for the Texas Weather cluster allowed some students—who knew the proper tools for measuring each phenomenon (e.g., wind speed)—to lose credit for correctly matching tools with phenomena. This occurred because, when these students ran the simulation, they simply manipulated the tools and overlooked the drop-down menu for choosing the phenomenon they intended to measure. The simulation ran as intended under these conditions, so there was nothing to cue the students that they were inadvertently losing points.
- Scoring rubrics should be reviewed to make sure that they are constructed in a consistent manner and conform to the task demands they are intended to measure. In the cognitive interviews, some rubrics awarded a point for meeting a single, straightforward criterion, while others required that the student do several things correctly. For example, in item 1 in the Galilean Moons cluster, students got 1 score point for each of the moons for which they correctly measured the maximum distance from Jupiter. On the other hand, in item 1 of the Redwall Limestone cluster, students had to correctly identify six different animals as being found, or not found, in Arizona to earn any credit.

We recommend that the second type of rubric (requiring students to do several things correctly) be limited to cases in which integration across knowledge is the construct of interest.

3.1.2 Cluster Score Distributions and Average Time to Complete, by Grade Level

Elementary School Clusters

As shown in Table 2, average time to complete the elementary school clusters ranged from six minutes for the Redwall Limestone cluster to 12 minutes for the Desert Plants cluster.

Table 2. Maximum Score and Average Time to Complete: Elementary School Clusters

Cluster Name	Maximum Score	Average Time to Complete
Desert Plants	9	12
German Pyramid Candle	4	9
Redwall Limestone	4	6
Terrarium Matter Cycle	9	11

Table 3 and Table 4 show the score distributions for elementary school clusters with maximum scores of four and nine, respectively.

The Redwall Limestone cluster was easy for all students, with 12 students (71%) earning three or 4 score points. Utah students did even better, with half earning the maximum score of four points and two others earning 3 points.

The Desert Plants cluster was also relatively easy, with 15 students (83%) earning at least four of the nine points possible. All six Utah students earned scores in this range. Further, two Utah students were the only ones who earned the maximum score of eight, and four of the five students who earned at least seven points were from Utah.

The Terrarium Matter Cycle cluster was harder for all students, with only four students (22%) earning at least four of the nine points possible. Half of the Utah students earned scores in this range. No student earned the full nine points on this cluster, but the highest scoring student was a Utah student who earned seven points.

The German Pyramid Candle was the hardest cluster, with only one student (from Utah) earning the maximum score of four points (and none earning 3 points). Further, seven students (41%) earned no credit, but only one Utah student was included in this group.

Table 3. Number of Students Attaining Cluster Total Scores in Specified Range: Elementary School Clusters with Maximum Score = 4

Cluster Name	Score 4–3	Score 2–1	Score 0
German Pyramid Candle	1	9	7
Redwall Limestone	12	4	1

Note. For both clusters, n = 17.

Table 4. Number of Students Attaining Cluster Total Scores in Specified Range: Elementary School Clusters with Maximum Score = 9

Cluster Name	Score 9–7	Score 6–4	Score 3–1	Score 0
Desert Plants	5	10	2	1
Terrarium Matter Cycle	1	3	13	1

Note. For both clusters, n = 18.

Middle School Clusters

As shown in Table 5, the average time to complete the middle school clusters ranged from 10 minutes for the Galilean Moons cluster to 14 minutes for the Texas Weather cluster.

Table 5. Maximum Score and Average Time to Complete: Middle School Clusters

Cluster Name	Maximum Score	Average Time to Complete
Galilean Moons	9	10
Hippos	10	10
Morning Fog	9	12
Texas Weather	11	14

Table 6 through Table 8 show the score distributions for middle school clusters with maximum scores of nine, 10, or, 11, respectively.

Students performed best on the Galilean Moons cluster with five students (42%) earning at least seven points and an additional four students (33%) earning between six and four points.

The Hippos cluster was also fairly easy, with seven students (58%) earning four or more points.

The Morning Fog and Texas Weather clusters (maximum scores nine and 11, respectively) were both challenging for students. Only five students (43%) earned scores greater than three on Morning Fog, and only four students (33%) earned scores greater than three on the Texas Weather cluster.

Table 6. Number of Students Attaining Cluster Total Sores in Specified Range: Middle School Clusters with Maximum Score = 9

Cluster Name	Score 9–7	Score 6–4	Score 3–1	Score 0
Galilean Moons	5	4	3	0
Morning Fog	2	3	7	0

Note. For both clusters, $n = 12$.

Table 7. Number of Students Attaining Cluster Total Scores in Specified Range: Middle School Clusters with Maximum Score = 10

Cluster Name	Score 10–7	Score 6–4	Score 3–1	Score 0
Hippos	2	5	3	0

Note. $n = 10$.

Table 8. Number of Students Attaining Cluster Total Scores in The Specified Range: Middle School Clusters with Maximum Score = 11

Cluster Name	Score 11–7	Score 6–4	Score 3–1	Score 0
Texas Weather	0	4	8	0

Note. n = 12.

High School Clusters

As shown in Table 9, the average time to complete the high school clusters ranged from 10 minutes for the Tuberculosis cluster to 19 minutes for the Blood Sugar Regulation cluster.

Table 9. Maximum Score and Average Time to Complete: High School Clusters

Cluster Name	Maximum Score	Average Time to Complete
Blood Sugar Regulation	7	19
Saving the Tuna	7	14
Tomcods	8	17
Tuberculosis	5	10

Table 10 through Table 12 show the score distributions for high school clusters with maximum scores of five, seven, or eight, respectively.

Students found all the high school clusters challenging but performed the worst on the Tomcods cluster. Only one student (7%) earned a score greater than three on this eight-point cluster, and four students (31%) earned no credit. Similarly, there were four students in both the Tuberculosis and Saving the Tuna clusters who earned no credit. No one earned more than 5 points on the seven-point Blood Sugar Regulation cluster, but scores for most students (9 out of 12) were solidly in the mid-range of 5 to 3 points.

Table 10. Number of Students Attaining Cluster Total Scores in Specified Range: High School Clusters with Maximum Score = 5

Cluster Name	Score 5–4	Score 3–1	Score 0
Tuberculosis	1	9	4

Note. n = 14.

Table 11. Number of Students Attaining Cluster Total Scores in Specified Range: High School Clusters with Maximum Score = 7

Cluster Name	Score 7–6	Score 5–3	Score 2–1	Score 0
Blood Sugar Regulation	0	9	3	1
Saving the Tuna	1	2	5	4

Note. Blood Pressure Regulation *n* = 13; Saving the Tuna *n* = 12.

Table 12. Number of Students Attaining Cluster Total Scores in Specified Range: High School Clusters with Maximum Score = 8

Cluster Name	Score 8–6	Score 5–4	Score 3–1	Score 0
Tomcods	0	1	9	4

Note. n = 14.

3.2 Detailed Discussion by Cluster: Elementary School

3.2.1 Cluster 1: Desert Plants

Performance Summary

The median time to complete the Desert Plants cluster was 11.5 minutes. Table 13 and Table 14 indicate the number of students attaining cluster total scores and item scores within the specified ranges, respectively.

Table 13. Number of Students Attaining Cluster Scores in Specified Range: Desert Plants

Score 9–7	Score 6–4	Score 3–1	Score 0
5	10	2	1

Note. Maximum score = 9; *n* = 18.

Table 14. Number of Students Attaining Item Scores in Specified Range, by Item: Desert Plants

	Maximum Item Score	Score 1	Score 0
Item 1 (Part A)	1	12	6
Item 1 (Part B)	1	13	5
Item 2 (Part B)	1	3	15

	Maximum Item Score	Score 3	Score 2–1	Score 0
Item 2 (Part A)	3	2	13	3
Item 3	3	14	3	1

Note. n = 18.

Students did relatively well on this cluster, but Item 2 was much more challenging than Items 1 or 3.

Task Demands

The following are task demands of the Desert Plants cluster:

- Organize or summarize data to highlight trends and patterns and/or determine relationships between the traits of an organism and survival in its environment.
- Understand and generate simple bar graphs or tables that document patterns, trends, or relationships between traits of an organism and its survival in a particular environment.

- Identify patterns or evidence in the data that support inferences about characteristics of an organism and those of its environment.
- Based on the provided data, identify or describe a claim regarding the relationship between the characteristics of an organism and survival in a particular environment.
- Evaluate the evidence to sort relevant from irrelevant information regarding survival of an organism in a particular environment.

Stimulus

The stimulus for the Desert Plants cluster is shown in Figure 1.

Figure 1. Stimulus: Desert Plants

Plant Survival in the Desert

Mesquite trees and cactus plants are both common in the Sonora Desert of North America, even though this region receives less than 15 inches of rain a year. In comparison, bird's nest ferns are common to the rainforests of southeastern Asia, where rainfall is often more than 100 inches a year.

These three plants have differences in their roots, stems, and leaves. The Characteristics of Plants table summarizes the characteristics of each type of plant.

Characteristics of Plants

	Mesquite Tree	Cactus Plant	Bird's Nest Fern
Roots	Long deep roots	Wide shallow roots	Short shallow roots
Stems	Non-expandable trunk	Thick expandable trunk	Thin stems
Leaves	Small leaves	Leaves reduced to thin spikes	Large leaves

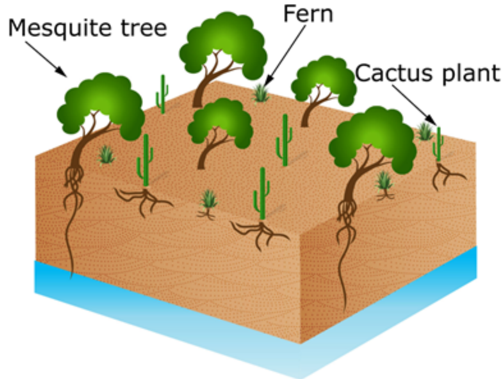
Plants use their roots, stems, and leaves to get and keep water. Differences in these structures affect the way in which different plants meet their needs for water.

Effect of Plant Structures on Ability to Get and Keep Water

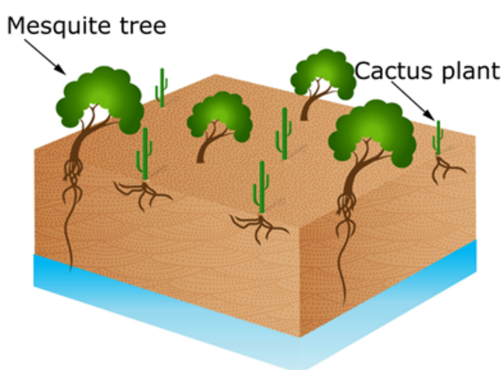
Plant Structure	Effect
Roots	Deep roots—allow plants to reach ground water below surface Wide shallow roots—allow plants to absorb a lot of water quickly when it rains
Leaves	Small waxy leaves—prevent loss of water in the hot sun
Stems	Thick expandable stems—allow plants to store water

To test how different characteristics affect a plant's ability to survive with less than 15 inches of rain a year, scientists planted Mesquite trees, cactus plants, and bird's nest ferns in a desert environment. A year later, they recorded how many of each type of plant survived.

Start of Experiment



End of Experiment



In the questions that follow you will construct an argument for why some plants survive better in the desert than others.

Details by Item

<u>Item 1</u>

Item 1 of the Desert Plants cluster is shown in Figure 2.

Figure 2. Item 1: Desert Plants

The following question has two parts. First, answer part A. Then, answer part B.

Use the data from the experiment to compare the survival of the three types of plants in the desert.

Part A

Record the data from the experiment by adding numbers to the table.

	Mesquite Trees	**Cactus Plants**	**Bird's Nest Ferns**
Number of plants at start of experiment			
Number of plants at end of experiment			

Part B

Select the **two** statements that are supported by the data in the table you created.

- ☐ All types of plants can survive in all environments.
- ☐ No types of plants can survive in a dry desert environment.
- ☐ All types of plants can survive in the dry desert environment.
- ☐ Some types of plants cannot survive in the dry desert environment.
- ☐ Some types of plants survive better than others in the dry desert environment.

<u>Item 1 (Part A)</u>

SCORES

Half of the California students (six) and all of the Utah students (six) earned credit (1 score point) on Part A.

COMPREHENSION

Those students who received credit for this item did not appear to be confused by any features of the item.

However, the students who did not receive credit seemed to have a general lack of comprehension of what was being asked. For example,

- one student wrote incoherent sentences instead of numbers;
- a second student decided to start at 27 "as a random number to start with"; and

- a third student said, "For mesquite trees, I got the start of experiment 1, do you see you start with 1, and at the end I saw how much they had altogether, and I got 3, so I was guessing that's how much it was." For the cactus plants, the student said, "I thought the same thing—they started off with 1 then ended with 3." For the bird's nest ferns, he said, "I was thinking the same thing because I was looking at the characteristics of plants—you start with 1 then you end with 3."

REASONING

The 12 students who earned credit all made sensible use of the experiment data.

> For example, one student said she counted the trees, plants, and ferns in the *Start of the Experiment* exhibit and began entering the numbers in the first row of the table. She explained, "I put 5 mesquite trees, because when I counted, there was 5 [at the beginning of the experiment]. When I counted the cactus, there was 5. And then the same for bird's nest ferns." She counted the trees, plants, and ferns in the *End of the Experiment* exhibit and began entering the numbers in the second row of the table. The student noted that there were four mesquite trees, explaining that this was "[b]ecause one of them had died during the experiment. And then for the cactus plants, the number stayed the same, at 5, because they normally live there, like, a lot, and they really don't need a lot of water to survive. And then the bird ferns all died during the experiment, so then that is a total of 0."

Item 1 (Part B)

SCORES

Thirteen students, including five of the six Utah students, earned credit (1 point) on Part B, which required them to identify two statements that are supported by the table in Part A. (One of these students did not receive credit for Part A but understood the general concept.)

COMPREHENSION

No features of this item appeared to confuse students.

REASONING

Most students used credible reasoning from evidence to reach a solution.

> For example, one student chose the second-to-last option (*[s]ome types of plants cannot survive in the dry desert environment*) because "at the start of the experiment, there was a total of five bird's nest ferns and then they all died, and also because one of the mesquite trees – they died – but I mean, most of them still remained." And she chose the last option (*[s]ome types of plants survive better than others in the dry desert environment*) because "out of all three of the plants, the cactus all lived instead of dying." She shared that she did not chose the first option (*[a]ll types of plants can survive in all environments*) because "As you can see, some of them died – like the bird's nest ferns and the mesquite trees." She shared that she did not choose the second option (*[n]o types of plants can survive in a dry desert environment*) "because the cactus – they still lived." She shared that she did not choose the third option (*[a]ll types of plants can survive in the dry desert environment*) "because the bird's nest ferns died."

Item 2

Item 2 of the Desert Plants cluster is shown in Figure 3.

Figure 3. Item 2: Desert Plants

The following question has two parts. First, answer part A. Then, answer part B.

Determine which traits of the three types of plants affect their survival in the desert.

Part A

The three tables show traits of each type of plant from the experiment. Select the boxes to identify whether each trait helps or does not help each plant survive in the desert.

Mesquite Tree Traits

	Helps Survival	Does Not Help Survival
Long deep roots	☐	☐
Non-expandable trunk	☐	☐
Small leaves	☐	☐

Cactus Plant Traits

	Helps Survival	Does Not Help Survival
Wide shallow roots	☐	☐
Thick stem	☐	☐
Thin spikes as leaves	☐	☐

Bird's Nest Fern Traits

	Helps Survival	Does Not Help Survival
Short shallow roots	☐	☐
Thin stem	☐	☐
Large leaves	☐	☐

Part B

Type a number into each box to identify the number of traits that help or do not help the plants survive, based on the tables in part A.

	Helps Survival	Does Not Help Survival
Mesquite Trees		
Cactus Plants		
Bird's Nest Ferns		

<u>Item 2 (Part A)</u>

SCORES

Points were awarded based on the number of plants for which the student correctly identified the traits that help the plant survive. Two students earned 3 score points (full credit) on Part A, six students earned 2 score points, and seven students earned 1 score point.

COMPREHENSION

No features of this item appeared to confuse students.

REASONING

Nine of the students used the *Characteristics of Plants* and *Effects of Plant Structures on Ability to Get and Keep Water* tables, and at least three of these students also referred to the exhibits showing plants that were alive at the beginning and end of the experiment. However, they did not necessarily interpret all the data correctly. For example, the following student referenced the information in the stimulus tables frequently and appropriately but misinterpreted some of the data. She did not appear to use the exhibits on the start and end of the experiment to check her understanding of which traits help or hinder survival.

- For the mesquite tree she said, "the mesquite tree has long deep roots and also has small leaves," and checked *Helps Survival* for roots and leaves. She continued, "The [mesquite] plant—I don't think that the non-expandable trunk will help. It says that thick expandable stems allow plants to store water, except the tree doesn't have one, so it can't store a lot of water, so I don't think that will help it survive." She checked *Does Not Help Survival* for the non-expandable trunk.

- For the cactus plant she said, "The cactus plant traits, it says it has wide shallow roots that allow the plant to absorb lots of water when it rains. So that would help it survive." She checked *Helps Survival* for roots. She continued, "The thick trunk also will, but thick stem would do that." She checked *Helps Survival* for trunk. She continued, "Then thin spikes as leaves—that probably wouldn't help them a lot." She checked *Does Not Help Survival* for leaves.

- For the bird's nest fern she said, "So for the bird's nest fern traits, it has shallow roots, and shallow roots allow it to absorb a lot of water when it rains, so that would probably help survive." She checked *Helps Survival* for roots. She continued, "A thin stem—that would probably not help it survive since the thin stem would not be able to hold a lot of water to help it survive." She checked *Does Not Help Survival* for the stem. She continued, "Then large leaves—that would probably be good. And small waxy leaves have lots of water in the hot sun. Yep." She checked *Helps Survival* for leaves.

Seven students made little or no use of the data in the stimulus and based their reasoning for Part A on prior knowledge or conjecture.

Item 2 (Part B)

SCORES

On Part B, most students quickly filled out the table on the number of traits that help or do not help each plant survive based on their responses in Part A.

However, only three students completed all six cells correctly, as required to earn credit (1 score point) on Part B.

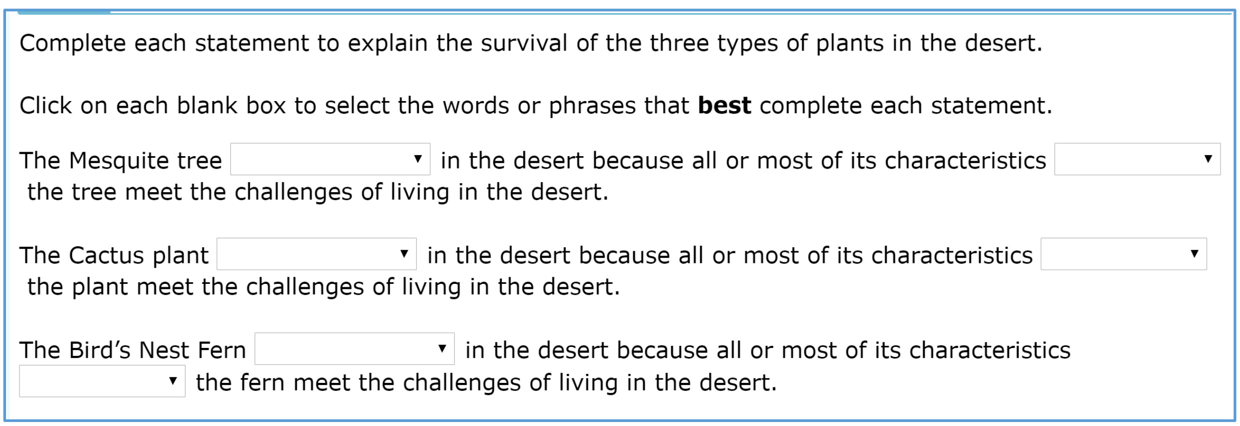
COMPREHENSION

On Part B, three students wrote the types of traits in the response fields (e.g., long deep roots) rather than the number of traits as indicated in the instructions. One student also wrote some extraneous text. One other student wrote text that was mostly incoherent.

Item 3

Item 3 of the Desert Plants cluster is shown in Figure 4.

Figure 4. Item 3: Desert Plants


Complete each statement to explain the survival of the three types of plants in the desert.

Click on each blank box to select the words or phrases that **best** complete each statement.

The Mesquite tree [▾] in the desert because all or most of its characteristics [▾] the tree meet the challenges of living in the desert.

The Cactus plant [▾] in the desert because all or most of its characteristics [▾] the plant meet the challenges of living in the desert.

The Bird's Nest Fern [▾] in the desert because all or most of its characteristics [▾] the fern meet the challenges of living in the desert.

SCORES

Students earned 1 point for each statement they completed correctly. Fourteen students completed all three statements correctly and earned full credit. This included all six of the Utah students.

Sixteen students earned a score point for the statement on the mesquite tree. Sixteen students earned a score point for the statement on the cactus plant, and 15 students earned a score point for the statement on the bird's nest fern.

COMPREHENSION

All students navigated through this item with ease.

REASONING

Most students used their answers to previous questions in the cluster to select responses from the drop-down menus. At least five students used information from the stimulus, and three students used prior knowledge.

The following is an example of a student who reasoned appropriately from the evidence in the stimulus to respond to Item 3:

> The student selected *survived well* for mesquite tree, explaining that this was “because all or most of its characteristics helped the tree meet the challenges of living in the desert; because the characteristics, such as having the long deep roots and the small leaves can help it survive in the desert.” She selected *survived best* for cactus plant, “because all or most of its characteristics helped it meet the challenges of living in the desert; because, of all of the plants, it stayed alive, and the characteristics such as having wide shallow roots and thick stems helped it live.” The student selected *did not survive* for bird’s nest fern, noting that “only one of its traits helped, and the rest—the two other ones—did not help it.” Then she selected the answers for the second part of each item, choosing *helped* for mesquite tree, *helped* for cactus plant, and *did not help* for bird’s nest fern.

3.2.2 Cluster 2: German Pyramid Candle

Performance Summary

The median time to complete the German Pyramid Candle cluster was nine minutes. Table 15 and Table 16 indicate the number of students attaining cluster total scores and item scores within the specified ranges, respectively.

Table 15. Number of Students Attaining Cluster Total Scores in Specified Range: German Pyramid Candle

Score 4–3	Score 2–1	Score 0
1	9	7

Note. Maximum score = 4. *n* = 17; one student ran out of time before attempting this cluster.

Table 16. Number of Students Attaining Item Scores in Specified Range, by Item: German Pyramid Candle

	Maximum Item Score	Score 2	Score 1	Score 0
Item 1	2	3	5	9

	Maximum Item Score	Score 1	Score 0
Item 2	1	2	15
Item 3	1	5	12

Note. n = 17; one student ran out of time before attempting this cluster.

This was the most difficult of the elementary school clusters; only one student (from Utah) earned full credit (4 points).

Task Demands

The following are task demands of the German Pyramid Candle cluster:

- Identify from a list, including distractors, the materials/tools needed for an investigation of how energy is transferred from place to place through heat, sound, light, or electric currents.
- Identify the outcome data that should be collected in an investigation of how energy is transferred from one place to another through heat, sound, light, or electric currents.
- Make and/or record observations about the transfer of energy from one place to another via heat, sound, light, or electric currents.
- Interpret and/or communicate the data from an investigation.

- Select, describe, or illustrate a prediction made by applying the findings from an investigation.

Stimulus

The stimulus for the German Pyramid Candle cluster is shown in Figure 5.

Figure 5. Stimulus: German Pyramid Candle

A German pyramid candle is a decoration whose parts only move when the candles are lit. The parts that move are driven by a fan that sits on the top of the pyramid. As the fan turns, other parts of the pyramid turn. The animation shows an example of a German pyramid candle. Click the small gray arrow to begin the animation.

Use the following questions to determine how energy is transferred from the candles to the fan blades.

Details by Item

Item 1

Item 1 of the German Pyramid Candle cluster is shown in Figure 6.

Figure 6. Item 1: German Pyramid Candle

In the following table, select the **two** pieces of data that explain how the candles affect the fan, and then use the animation to describe the relationship between these two variables.

Relationship of Outcome Data

Variables	Relationship
▾	▾
▾	▾

SCORES

Two (Utah) students earned full credit (2 score points) on this item, which required students to identify two variables that explain the influence of the candles on the fan and then describe the relationship between these variables.

Seven other students earned partial credit for selecting the two correct variables but not correctly specifying the relationships—five were Utah students.

Additional students selected at least one of the correct variables.

A total of 13 students correctly selected the temperature of the air between the blades and the candles as one of the variables, and eight students correctly selected the rotation speed of the blade.

COMPREHENSION

Students clearly did not understand how to describe the relationship between the two variables as only four students entered any responses to this part of the question. It is not clear how much of the confusion was because the students did not understand how energy was transferred and how much of the confusion was due to not understanding what the question was asking.

Five students were hesitant about the entire item, and two students tried to guess at the relationships between the two variables because they did not really understand what "the relationship" meant.

REASONING

Most students tried to reason their way to a solution but lacked the content knowledge to do so without error. The following shows the reasoning process for one student who exemplifies this:

> The student said, "The first variable is probably going to be *brightness* because if they're more brighter, it probably means that it's hotter. And for relationship, I'm going to do *increase* because I think it turns because something is taking in the heat energy and it's using the heat energy from the candles to rotate the fan, and that's why the brightness of the candles would probably increase the speed of the rotation of the fans. And so for variable two, I'm going to do the *temperature of the air between the blades and the candles* – I chose that because if the air is colder or cooler, it's probably not going to rotate that much because it takes in the heat energy that the candles create and it rotates them . . . And if it's like hot or warm, it's probably going to rotate faster . . . if I'm correct. And for the relationship, I'm going to do decrease because if it's slower or cooler, it's probably going to be less . . . or not as fast as if it was warmer."

Item 2

Item 2 of the German Pyramid Candle cluster is shown in Figure 7.

Figure 7. Item 2: German Pyramid Candle

Use the table below to correctly order the statements based on what you have observed. Use the numbers 1 through 4 to order your statements, 1 being the first step and 4 being the last step. Use the "-" sign to indicate that the statement is not a part of the process you observed.

Step	Statement
▾	Air moves upward past the fan blades
▾	Light from candles transfers energy to the air
▾	Air gets hotter
▾	Moving air transfers energy to the fan blades
▾	Air transfers heat energy to the fan blades
▾	Heat from candles transfers energy to the air
▾	Light energy carries the air upwards past the fan blade

Scores

All but one student observed the whole animation, but only two (Utah) students earned credit (1 score point) on this item by correctly ordering the steps based on what they observed in the animation.

Comprehension

One student did not seem to understand that he was to order the steps, and it was not clear how he selected the numbers for his responses.

Reasoning

Students had the same issues with lack of content knowledge as they did with Item 1.

> For example, one student correctly chose *[h]eat from candles transfers energy to the air* for step 1 (noting that "the energy carries the air upward past the fan"), but faltered after that. She chose *[a]ir transfers heat energy to the blades* for step 2, noting that it "was going to the fan blades." For step 3, the student initially chose *[a]ir moves upward past the fan blades* but changed it to *[l]ight energy carries the air upwards past the fan blade*. When prompted later to explain why she changed her answer, she explained, "Because it made more sense if hot air moved upward past the fan blades, but it was just air, so I was thinking light energy carries the air upward past the fan blades because first the energy goes to the fan blades and then the light energy from the candles goes past the fans." For step 4, she thought for a moment and said, "I think this (*air gets hotter*), and chose it," explaining "because it goes around more."

Item 3

Item 3 of the German Pyramid Candle cluster is shown in Figure 8.

Figure 8. Item 3: German Pyramid Candle

With your knowledge of the process that drives the German pyramid candle, select the boxes in the table to indicate whether or not the changes listed would affect the animation.

	Affect	Not Affect
Change the number of candles	☐	☐
Remove the air from between the candles and the blades	☐	☐
Change the amount of wax on the candles	☐	☐
Change the angle of the blades	☐	☐
Change the color of the fan blades	☐	☐

SCORES

Five students earned credit (1 score point) for this item.

Nine other students correctly classified four of the five changes, but earned no credit, based on the scoring rubric.

COMPREHENSION

No features of this item appeared to confuse students.

REASONING

As with the other items in this cluster, students needed prior content knowledge to reason their way to a correct solution. For example, one student, who had most of the requisite knowledge, said,

> "For the first one, the *change in number of candles*, I think that, with more heat and light, I think it will affect it a little bit more by making the blades spin faster. *Removing the air from between the candle and blades*, I think that will affect it because the GPC probably takes in the air from what's underneath it. For the third one, the *change in the amount of wax on the candles*, I think that will not affect it because the wax just increases the duration of the candle, which wouldn't affect it. *Change the angle of the blades*, I don't think that would affect it because if you just turn the blades over to at least an angle where it looks like it's even, I don't think that will affect it either. *Change the color of the fan blades*, I don't think changing the color of the fan blades would affect it because it's just color, and it's for decoration most of the time."

3.2.3 Cluster 3: Redwall Limestone

Performance Summary

The median time to complete the Redwall Limestone cluster was six minutes. Table 17 and Table 18 indicate the number of students attaining cluster total scores and item scores within the specified ranges, respectively.

Table 17. Number of Students Attaining Cluster Total Scores in Specified Range: Redwall Limestone

Score 4–3	Score 2–1	Score 0
12	4	1

Note. Maximum score = 4; *n* = 17; one student ran out of time before attempting this cluster.

Table 18. Number of Students Attaining Item Score in Specified Range, by Item: Redwall Limestone

	Score 1	Score 0
Item 1	13	4
Item 2	13	4
Item 3 (Part A)	14	3
Item 3 (Part B)	7	10

Note. Maximum score for each item = 1; *n* = 17; one student ran out of time before attempting this cluster.

Task Demands

The following are task demands of the Redwall Limestone cluster:

- Organize or summarize data to highlight trends, patterns, or correlations between plant and animal fossils and the environments in which they lived.
- Generate graphs or tables that document patterns, trends, or correlations in the fossil record.
- Identify evidence in the data that support inferences about plant and animal fossils and the environments in which they lived.

Stimulus

The stimulus for the Redwall Limestone cluster is shown in Figure 9.

Figure 9. Stimulus: Redwall Limestone

The Grand Canyon is a huge canyon located in Arizona. The canyon has been formed by the Colorado River. The river has cut down into the ground, exposing rock layers that were deposited millions of years ago. The picture shows part of the Grand Canyon.

Portion of Grand Canyon

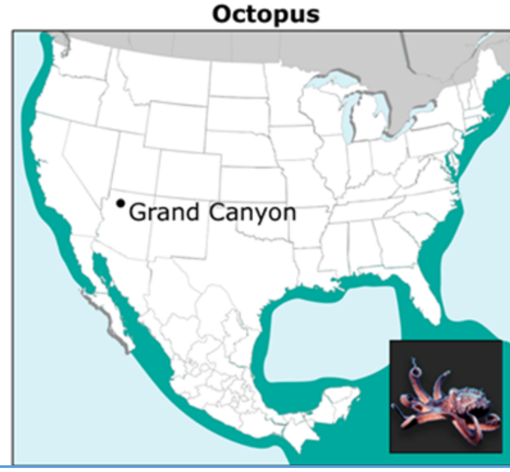
One of these rock layers is called the Redwall Limestone. The Redwall Limestone contains many different fossils, including corals, clams, octopi, and fish.

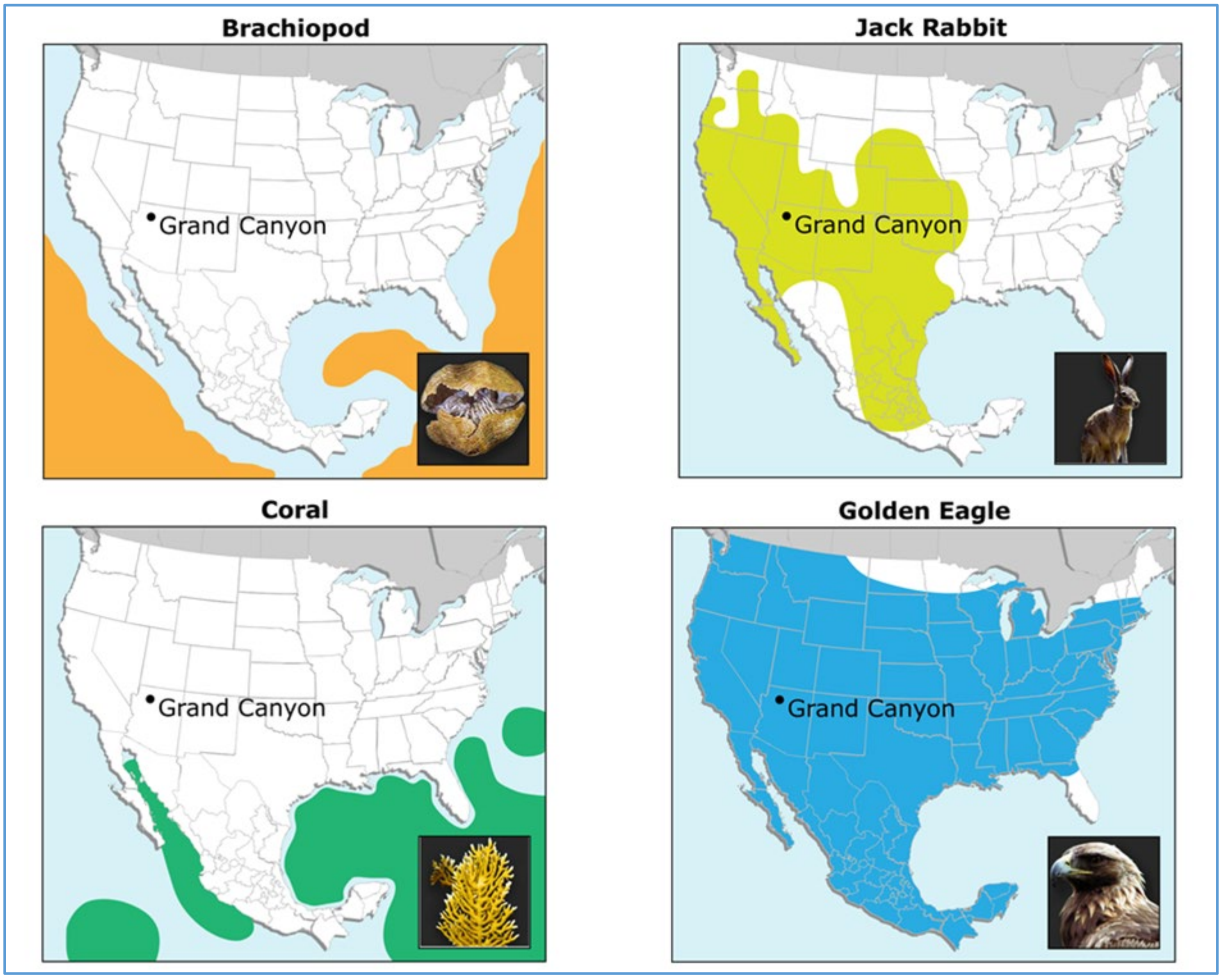
In the questions that follow, you will study six animals in order to lean about what Arizona was like when the Redwall Limetone was deposited millions of years ago.

The pictures show the animals and maps of where they are found. The colored regions show where the animals live.

Bighorn Sheep

Octopus





Despite some incorrect responses, nearly all the students seemed comfortable navigating through the maps to decide where the animals are found and filling out the tables in Items 1 and 2. One student did not make any use of the maps.

Details by Item

Item 1

Item 1 of the Redwall Limestone cluster is shown in Figure 10.

Figure 10. Item 1: Redwall Limestone

Using the given maps, complete the table by identifying whether each animal is found in Arizona.

	Found in Arizona	Not Found in Arizona
Bighorn Sheep	☐	☐
Octopus	☐	☐
Brachiopod	☐	☐
Jack Rabbit	☐	☐
Coral	☐	☐
Golden Eagle	☐	☐

Scores

Thirteen students earned credit (1 score point) on this item.

Comprehension

No features of this item appeared to confuse students.

Reasoning

Ten of the 13 students who earned credit showed evidence in the think-aloud of using the maps to reason their way to a solution, as intended.

For example, one student

- selected *Found in Arizona* for bighorn sheep "because the map that it gives you shows you that it's located in Arizona."
- selected *Not Found in Arizona* for octopus, explaining that "It's found in oceans – not really in the state."
- selected *Not Found in Arizona* for brachiopod, noting, with a laugh, "Because it's in the oceans, not the state – like the octopus . . . octopi."
- selected *Found in Arizona* for jack rabbit "because the map that it gives you shows it's located in Arizona."
- selected *Not Found in Arizona* for coral because "the map that it gives you has those green things that shows you that it's not located in Arizona."
- selected *Found in Arizona* for the golden eagle, noting that "the blue is all over the United States, so yeah, it's in Arizona."

Among the four students who did not earn credit for this item, each mis-located two of the six animals. The think-alouds showed that three of these students formed their answers based on background knowledge and some educated guessing rather than using the maps.

For example, one student

- selected *Not Found in Arizona* for bighorn sheep because "When I went to Arizona, I've never seen a bighorn sheep over there, so I really think it is not in there."
- selected *Found in Arizona* for jack rabbit, explaining that "it's in there because I've seen one when I went to Arizona."
- selected *Not Found in Arizona* for coral. This choice appeared to be at random, marked after the student said, "I've never heard of that animal too because in school we don't really learn about coral and so yeah I've never heard of it and I don't know if they're ever in Arizona, so . . ."
- selected *found in Arizona* for golden eagle because "I think it's in Arizona because our school mascot is the golden eagle and they always say golden eagles are from Arizona."

Item 2

Item 2 of the Redwall Limestone cluster is shown in Figure 11.

Figure 11. Item 2: Redwall Limestone

Using the given maps, complete the table by selecting whether each animal lives on land or in water.

Animal	Environment
Bighorn Sheep	▾
Octopus	▾
Brachiopod	▾
Jack Rabbit	▾
Coral	▾
Golden Eagle	▾

SCORES

Thirteen students earned credit (1 score point) on this item.

COMPREHENSION

No features of this item appeared to confuse students. All students worked through the item fairly quickly, and three of the students commented that it was easy.

REASONING

Among the 13 students who earned credit, most did not appear to make much use of the maps in formulating their responses, apparently because they felt that they could easily respond based on background knowledge about the animals.

> For example, one student shared that she knows bighorn sheep live on land and that octopi are living in the water. But then she noted that she wasn't sure about coral, adding, "Sometimes you see coral on the beach or somewhere else, and so I don't know if it's land or water. But maybe it was washed up on the beach, so I was thinking water."

Students who did not earn credit for this item mis-located either the brachiopod or the coral; one student also mis-located the golden eagle. These students also relied on background knowledge for their responses. For example, one student explained his choices as follows:

- The bighorn sheep "is on land because I don't think he'll make it in the water."
- The octopus "has to live in the water to survive."
- The brachiopod "has to live in the water because it looks like a jellyfish and jellyfishes have to live in the water, so I thought maybe that does too, and I looked at the picture and thought it has to live in the water."
- "I looked at [the jack rabbit], and that's a land animal, and regular rabbits live on land, and that's why I picked that one."
- "[The coral] has to be on land because it kind of looks like a tree and trees have to be on land."
- "Birds and eagles are on land, so I picked that eagle to be on land, so I just knew it from my knowledge."

Item 3

Item 3 of the Redwall Limestone cluster is shown in Figure 12.

Figure 12. Item 3: Redwall Limestone

The following question has two parts. First, answer part A. Then, answer part B.

Part A

Within the Grand Canyon, a rock layer contains fossils of octopi (plural of "octopus"), brachiopods, and corals. What can you conclude about the environment of the Grand Canyon region from the fossil evidence?

Ⓐ The Grand Canyon region was always desert.

Ⓑ The Grand Canyon region was once underwater.

Ⓒ The Grand Canyon region experienced a lot of rain.

Ⓓ The fossils do not provide any information about the environment.

Part B

Which statement supports your conclusion?

Ⓐ The rock layer contains fossils of only animals that live in water.

Ⓑ The rock layer contains fossils of only animals that live on land.

Ⓒ The rock layer contains fossils of animals that live neither on land nor in water.

Ⓓ The rock layer contains fossils of animals that live on land and animals that live in water.

Item 3 (Part A)

SCORES

Fourteen students earned credit (1 score point) on this sub-item.

There was no common theme to the wrong answers—there were three possible wrong answers, and each of the three students who failed to earn credit chose a different one.

COMPREHENSION

Among the three students who did not earn full credit for the sub-item, one student appeared not to understand what the question was asking. She said she was confused on how to respond because "I thought it was going to ask me 'does it usually rain there?' and it doesn't usually rain there because it's in Arizona."

REASONING

The 14 students who earned credit for this sub-item (1 score point) all appeared to evaluate the possible response option against credible criteria as they reasoned their way to a solution.

> For example, one student first read option A, *[t]he Grand Canyon region was always desert*, out loud. Then he said he wanted to check the next option and read *[t]he Grand Canyon region was once underwater*. The student said that option B could be the answer, "but the first option [A] is not because it said in the question [the fossils] were sea animals." The student then read option C, *[t]he Grand Canyon region experienced a lot of rain*, and option D, *[t]he fossils do not provide any information about the environment*. He said that it can't be option D because "[the question] doesn't have anything to do with the animals that are living today." He said it probably wasn't option C because "even if it rained, [but] it wasn't an ocean, then the coral couldn't live there." The student concluded that the correct answer had to be B.

Item 3 (Part B)

Scores

Seven students earned credit (1 score point) on this sub-item.

Comprehension

Among the 10 students who did not earn credit on this sub-item, most appeared to be confused as to what the question was asking. Rather than associating the question with Part A, these students appeared to be trying to answer a separate question about the types of animal fossils that might be found in the canyon walls. Further, they did not seem to know where to look for information that would help them answer the question; they tended to reference the list of *current-day* animals mentioned in the stimulus, and to do so irrespective of whether these animals were found in Arizona. Consequently, nine of these 10 students selected option D, *[t]he rock layer contains fossils of animals that live on land and animals that live in water*, using reasoning such as the following:

> One student said, "obviously C, *the rock layer contains fossils of animals that live neither on land nor in water*, is wrong, it's not only water because they have jack rabbits, the goat-ram thing, and the eagle so that's not true." For option B, *the rock layer contains fossils of only animals that live on land*," he said: "that's not true, there are octopus, coral and brachiopod." He read out loud response option C a second time, *the rock layer contains fossils of animals that live neither on land nor in water*, and said "the bird does live on land and it flies a lot, but it's still on land, so it has to be D, *the rock layer contains fossils of animals that live on land and animals that live in water*."

Some students also seemed to have problems with the structure of the answer choices (A, or B, or neither A nor B, or both A and B).

> For example, one student said, "What I found confusing was this one since I was looking at D and it said, 'live in water' at the end, just like A, so I was looking at it, and I figured out that it said lived on land AND on water. It kind of confused me just looking at the end that both of them said 'live in water.'"

Reasoning

The seven students who earned credit for this sub-item all appeared to use credible criteria in reasoning their way to a solution.

For example, one student read out loud the stem and option A, *[t]he rock layer contains fossils of only animals that live in water*. He said that it could be that one, but he wanted to read the other options. He read out loud option B, *[t]he rock layer contains fossils of only animals that live on land*. The student said, “no, it wouldn’t be that one because the answer [to Part A] doesn’t have anything to do with that.” He read option C, *[t]he rock layer contains fossils of animals that live neither on land nor in water*, and said it couldn’t be the right answer, because the question says that [the rock layer] has sea animals. He read option D*, [t]he rock layer contains fossils of animals that live on land and animals that live in water*. The student said that “the question never said anything about that part” and chose A.

3.2.4 Cluster 4: Terrarium Matter Cycle

Performance Summary

The median time to complete the Terrarium Matter Cycle cluster was 11 minutes. Table 19 and Table 20 indicate the number of students attaining cluster total scores and item scores within the specified ranges, respectively.

Table 19. Number of Students Attaining Cluster Total Scores in Specified Range: Terrarium Matter Cycle

Score 9–7	Score 6–4	Score 3–1	Score 0
1	3	13	1

Note. Maximum score = 9; *n* = 18.

Table 20. Number of Students Attaining Item Scores in Specified Range, by Item: Terrarium Matter Cycle

	Maximum Item Score	Score 1	Score 0
Item 1 (Part A)	1	3	15
Item 1 (Part B)	1	6	12
Item 2 (Part A)	1	8	7
Item 2 (Part C)	1	1	17
Item 2 (Part D)	1	1	17
Item 3	1	7	11

	Maximum Item Score	Score 3	Score 2–1	Score 0
Item 2 (Part B)	3	3	10	5

Note. n = 18

Earning credits on this cluster was challenging for the students. Two of the Utah students earned the most credit (seven and six credits respectively), likely reflecting their greater exposure to NGSS-based instruction.

Task Demands

The following are task demands of the Terrarium Matter Cycle cluster:

- Select or identify from a collection of potential model components, including distractors, the parts of a model needed to describe the movement of matter among plants, animals, decomposers, and the environment.

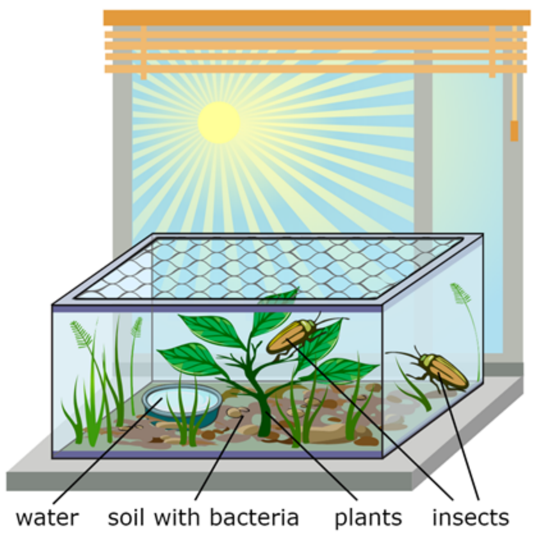
- Manipulate the components of a model to demonstrate properties, processes, and/or events that result in the movement of matter among plants, animals, decomposers, and the environment including the relationships of organisms and/or the cycle(s) of matter and/or energy.
- Articulate, describe, illustrate, select, or identify the relationships among components of a model that describe the movement of matter among plants, animals, decomposers, and the environment.
- Make predictions about the effects of changes in model components including the substitution, elimination, or addition of matter and/or an organism and the result.

Stimulus

The stimulus for the Terrarium Matter Cycle cluster is shown in Figure 13.

Figure 13. Stimulus: Terrarium Matter Cycle

A science class sets up four terrariums on a sunny windowsill. Each terrarium contains water and insects. Each one also contains a combination of gravel, soil with bacteria, and/or plants according to the Terrarium Setups table.



Terrarium Setups

	Terrarium 1	Terrarium 2	Terrarium 3	Terrarium 4
Soil			X	X
Gravel	X	X		
Plants		X		X

The students observe the terrariums every 5 days for 15 total days and record observations of the insects and plants. Their data are shown in the Terrarium Observations diagrams.

Terrarium 1 Observations

Day	Insects
1	Alive
5	Not alive
10	Not alive
15	Not alive

Terrarium 2 Observations

Day	Insects	Plants
1	Alive	Alive
5	Alive	Alive
10	Alive	Not alive
15	Not alive	Not alive

Terrarium 3 Observations

Day	Insects
1	Alive
5	Not alive
10	Not alive
15	Not alive

Terrarium 4 Observations

Day	Insects	Plants
1	Alive	Alive
5	Alive	Alive
10	Alive	Alive
15	Alive	Alive

In the following questions, you will develop a model to show why the insects only survive under certain environmental conditions.

Details by Item

Item 1

Item 1 of the Terrarium Matter Cycle cluster is shown in Figure 14.

Figure 14. Item 1: Terrarium Matter Cycle

The following question has two parts. First, answer part A. Then, answer part B.

Part A

Based on the observations of the terrariums, identify the parts that must be present for the insects to survive.

	Must be present
Gravel	☐
Soil with Bacteria	☐
Water	☐
Insects	☐
Plants	☐

Part B

Select the **three** statements that explain why these parts are necessary for the insects to survive.

- ☐ Insects need plants for food.
- ☐ Insects need soil to lay their eggs in.
- ☐ Plants need nutrients from the soil.
- ☐ Gravel is necessary for water drainage.
- ☐ Water is necessary for all living organisms.
- ☐ All living organisms take in matter from the environment.
- ☐ Different types of organisms are necessary for stable ecosystems.

Item 1 (Part A)

SCORES

Three students earned credit (1 score point) on this sub-item, which required them to correctly identify all four of the elements that must be present for the insects to survive. Ten other students correctly identified three of the four parts.

COMPREHENSION

Several students had trouble with the concept that the organism itself (i.e., insects) was one of the things that had to be present for that organism to survive. Six students gave a response that correctly identified soil with bacteria, water, and light as essential, but left out insects. Some others chose insects, but interpreted it as other insects, or were not sure.

> For example, when the interviewer asked after the think-aloud, “You weren’t sure whether to click insects or not here. Could you tell me a little about that?” One student said, “Yeah. Would it be the insects themselves? Or would it be different insects? Like you’d put two cockroaches in there with a ladybug. Or you’d put two ladybugs with a spider. I don’t know. If insects have to be there to survive, then yes, but if it is different insects and they’d be harmless, then I’d say no, they don’t need to be there. So maybe more description there.”

REASONING

The three students who received credit for the sub-item displayed the type of reasoning from evidence that was expected, although their reasoning was not necessarily correct in every detail.

> For example, one student said, “I know a class sets up four terrariums by a sunny windowsill, so light can get in to help the plants. I know plants have a photosynthesis process, and they need the sun to make food. There are also insects so they can eat, and water so they can drink, and soil so they can have a stable root because I know that plants don’t need soil to grow. In terrarium 3 and 4 there is soil, and in terrarium 1 and 2 there is gravel, and in 2 and 4 there are plants. A student observes the terrarium every 5 days for 15 days and records observation. Three times he observes them to collect observation—like the two living things in there, like the insects and the plants, and the data is shown on the diagram. I can see that the day 1 the insects are alive because in terrarium 1 there is only gravel, but no plants, so they don’t have anything to eat, so they can only survive about a day. Day 1, the insects are alive because—they are alive for three checks because they have gravel and plants . . . . The plants dying would probably be because maybe gravel is not strong to hold their roots. If the plants die, so do the insects. In terrarium 3, the insects are alive, and they all die on the next days because they don’t have any plants to eat. And then terrarium 4 has plants and soil, so it has plenty for the insects to eat, and it is a good support for the plants, so if they both stay alive, they can feed off each other.”

Many students who did not receive credit made only limited use of the experimental data provided in the stimulus and relied entirely or primarily on background knowledge.

> For example, for *Gravel*, one student said, “I don’t think it should be present because, if you just need gravel, you would have nothing to do with the soil in there.” For *Soil with Bacteria* the student said, “It must be present because a lot of plants and flowers, they need soil—and they also have bacteria in it or something.” For *Water*, the student said, “It definitely needs to be present because with just sun and soil, it won’t let it grow because every plant needs water, soil, and sun.” For *Insects*, the student said, “Yeah, because bees like going on sunflowers, so yeah it could be present.” For *Plants*, the student said, “Not so much cause if you’re going to grow one it’s already present . . . .” When asked if this was from the student’s prior knowledge, she agreed.

Item 1 (Part B)

Scores

Six students earned credit (1 score point) on this sub-item, which required students to correctly identify all three of the statements that explained why the elements in Part A are necessary for the insects to survive. Ten other students correctly identified two of the three statements.

Comprehension

No features of this item appeared to confuse students.

Reasoning

Students reasoned from background knowledge, but not necessarily content area knowledge gained in school.

> For example, one student selected option 1, and when asked how she knew, the student said, "if insects don't have food or water they'll die, and I know that just from background knowledge." The student selected option 3 because, "plants need nutrients from the soil, or they will die too… I just used my background knowledge." Student selected option 4 (*[g]ravel is necessary for water drainage*) and when asked how she knew, she said, "Just from learning it in school, I've just heard it before."

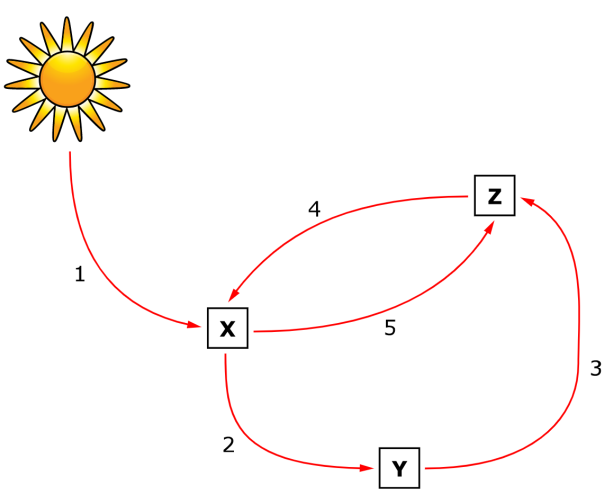
Item 2

Item 2 of the Terrarium Matter Cycle cluster is shown in Figure 15.

Figure 15. Item 2: Terrarium Matter Cycle

The Terrarium Cycle of Matter and Energy diagram shows an incomplete model of the terrarium environment.

Terrarium Cycle of Matter and Energy



The following question has four parts. First, answer part A. Next, answer part B. Then, answer part C. Finally, answer part D.

Part A

Select the boxes to identify X, Y, and Z.

	X	Y	Z
Gravel	☐	☐	☐
Soil with Bacteria	☐	☐	☐
Water	☐	☐	☐
Insects	☐	☐	☐
Plants	☐	☐	☐

Part B

Select the boxes to identify X, Y, and Z as a producer, consumer, or decomposer.

X: [▾]
Y: [▾]
Z: [▾]

Part C

Select the **two** numbers that represent arrows in the model to show when matter or energy is moved from the environment to organisms.

☐ 1

☐ 2

☐ 3

☐ 4

☐ 5

Part D

Carbon dioxide and water are missing from this model. If added, where would the arrow be pointing?

Ⓐ from X toward Y

Ⓑ from Y toward Z

Ⓒ from the environment toward X

Ⓓ from the environment toward Z

Students generally did not understand the *Terrarium Cycle of Matter and Energy* diagram in Item 2. One student did not answer any of the parts in Item 2.

Item 2 (Part A)

Scores

Only three students earned full credit (3 score points) on Part A, which required selecting correct labels for X, Y, and Z. Ten other students earned 1 score point. Two of the three students who earned full credit were from Utah.

Comprehension

Six students said Part A was confusing. They appeared not to understand the conventions of the diagram and possibly also did not understand the concept of matter and energy cycle.

> For example, one student said, "I don't get this question . . . I think it's missing something—the soil, the water, and insects that give it nutrients or something." The student attempted to click the diagram, thinking it might be interactive. She then moved on to Part A, read it aloud, and said, "I think for number 1 it's sun, then X is going to be *water*, and then this is going to be *insects*, and then this is going to be *plants*." After checking X for *Water*, the student also checked X for *Insects* and X for *Plants*. She then realized that she had overwritten her response to X twice and went back to check X for *Water*, Y for *Insects*, and Z for *Plants*.

Only one of the Utah students thought this sub-item was confusing; the remaining five Utah students did not express confusion or appear to guess at the interpretation of the diagram.

Item 2 (Part B)

Scores

Eight students earned credit (1 score point) in Part B by correctly identifying X, Y, and Z as a producer, consumer, or decomposer. Seven other students identified one of the components correctly.

Comprehension

Only one student expressed confusion on Part B, and this appeared to relate more to confusion over the producer, consumer, and decomposer roles than to the wording of the item. The student said:

> "What was confusing on this was B, because I forgot which one was that, so I was looking, and I thought about what was a producer, and I remembered that [it] was something that helps it grow. And X was the soil and bacteria, so X would have been the producer. The consumer got me confused because I didn't remember learning about the consumer. So, I was thinking it probably was the plants since I knew the decomposer was the one who would help the things decompose into the ground, and that was probably the insects. So, I knew that Y was the consumer."

REASONING

The reasoning of students who received credit for Part B indicated that they did know the facts of the natter and energy cycle, whether or not they understood the letters in the response choices as referencing the diagram.

> For example, one student said, "X is a *producer*, Y is a *consumer*, and Z has to be *decomposer* . . . X is producer because sunlight goes to the plants, and then the plants produce food for themselves and others, Y is consumer because the consumer eats the producer, and Z is decomposer, because after the consumer dies, the decomposer decomposes it and turns it into soil."

<u>Item 2 (Part C)</u>

SCORES

Only one (Utah) student earned credit (1 score point) on Part C, which required that students select both the arrows in the model that showed where matter or energy is moved from the environment to organisms. Nine other students correctly selected the arrow from the sun to X, but not the arrow from Z to X.

COMPREHENSION

The vocabulary used in this sub-item, particularly "environment," "organism," and "matter," was unfamiliar to several of the students.

> For example, one student did not understand the term "matter." The student said he was confused by "questions that had things to do with 'matter' because I know what matter is, but we started learning in science class, and I haven't fully gotten the sense of matter yet."

Confusion may also have arisen from the way in which the term "environment" is used, namely, to refer to the inanimate environment only.

REASONING

Most students tried to reason their way to a solution, but their content knowledge was too limited to allow them to identify both correct arrows. For example:

> One student said, "I'm going to say one of my answers is '1' because of light energy maybe is being moved from the environment, from the sun – I'm pretty sure that's part of the environment, and I'm pretty sure a plant is an organism. And for my second number I'm trying to think about what I can say . . . because the plant has matter, I'm pretty sure, or everything has matter. And a plant is an organism, and it says matter or energy, and the matter is being given or moved from the plant to the insect."

> Another student said, "I chose 2 and 3 since those are the necessary parts since the soil went in a circle to the soil. From the soil to the plants and from the plant to the insect. Since I thought that was the most important part. If it was 4 and 2, it would just be the same thing, but I thought 2 and 3 would be better and make more sense since the insect would be going to the soil and then the soil would make the plants and that wouldn't really make sense." The interviewer asks the student, "What do you think the question is asking?" The student

said, “It is showing that energy is moved from the environment to the organisms and I chose those since the matter in the sun is giving the soil energy to make the plants grow and that would keep going around. The plants would be decomposed or eaten by the bugs.”

Item 2 (Part D)

SCORES

Only three students earned credit (1 score point) on Part D, which asked where the arrow would be pointed if carbon dioxide and water were added to the model. Interestingly, eight students incorrectly indicated that the arrow would point from X toward Y.

COMPREHENSION

Several students simply lacked the content knowledge to answer this question.

> For example, one student said, "because I had to find from X toward Y – I had to know that the insects carried the carbon dioxide to the plants, but then also carry it to the soil."

Item 3

Item 3 of the Terrarium Matter Cycle cluster is shown in Figure 16.

Figure 16. Item 3: Terrarium Matter Cycle

Complete the table to identify your expected observations of the plants in a terrarium with only water, soil, and plants.

Day	Plants
1	▾
5	▾
10	▾
15	▾

SCORES

Seven students earned credit (1 score point) on this item.

COMPREHENSION

No issues with comprehension of the item were noted.

REASONING

Some students applied the information provided in the experiment to help them answer this question, although not all students were able to interpret the information from the experiment correctly.

> An example of using the experimental information correctly was a student who said, "This question is asking me to see how the plants, what I would observe if the plants were in a terrarium with water, soil, and plants. Plants would be plants, and soil would be soil, and water would be something to keep the plants alive. So, day 1 they would probably be alive.

> After 5 days, as long as plants are supplied by water and sun, they'd be alive. On day 10, they'd probably still be alive because of the ecosystem in the terrarium. On day 15, they could really be either, but I think that this question wants you to say, if they have everything they need, they'd be alive." After completing the cluster, when the interviewer asked the student if he used any information from the left side of the screen, the student said, "I used a lot of information from the left side of the screen because in terrarium 4 they stayed alive for 15 whole days, and just having soil, plants and water was not on that chart, but I bet they had it. I thought, since they stayed alive on that one, they'd stay alive in this one."
>
> Another student used the data from the terrarium experiment but without seeming to comprehend how to interpret the data. He said, "What I found confusing was on [day] 5 that [the terraria] were tied, and that 2 of them were alive and 2 of them were not alive. So that made it really confusing since I didn't know which one to choose."

At least 10 students, however, including some of those who earned credit, used only their prior content knowledge and/or personal experience to respond.

> For example, one student said, "Day 1: *alive*. I think I'll put *alive*. My plants have been alive for 2 weeks." She clicked *Alive* for days 1, 5, and 10. "*Alive*. I don't know if they're going to be alive so I'm going to try *Not Alive* (clicked *Not Alive* for day 15), I don't know. I've had tomatoes that lasted like months and months."

3.3 Detailed Discussion by Cluster: Middle School

3.3.1 Cluster 1: Galilean Moons

Performance Summary

The median time to complete the Galilean Moons cluster was 10 minutes. Table 21 and Table 22 indicate the number of students attaining cluster total scores and item scores within the specified ranges, respectively.

Table 21. Number of Students Attaining Cluster Total Scores in Specified Range: Galilean Moons

Score 9–7	Score 6–4	Score 3–1	Score 0
5	4	3	0

Note. Maximum score = 9; *n* = 12.

Table 22. Number of Students Attaining Item Scores in Specified Range, by Item: Galilean Moons

	Maximum Item Score	Score 4–3	Score 2–1	Score 0
Item 1	4	7	1	4
Item 2	4	7	4	1

	Maximum Item Score	Score 1	Score 0
Item 3	1	3	9

Note. n = 12.

Task Demands

The following are task demands of the Galilean Moons cluster:

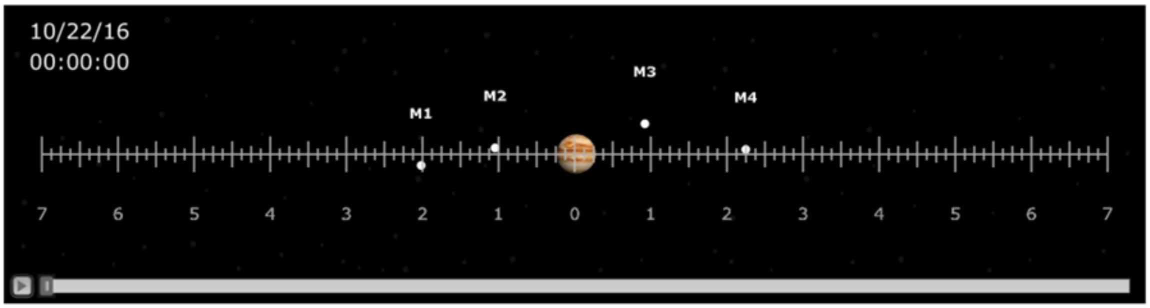
- Make simple calculations using given data to estimate the properties (e.g., mass, surface temperature, diameter) and locations of different solar system objects relative to a given reference point/object (Item 1).
- Calculate or estimate or identify properties of objects or relationships among objects in the solar system, based on data from one or more sources (Item 2).
- Given a partial model of objects in the solar system, identify objects or relationships that can be represented in the model or the reasons why they cannot be represented in the model (Item 3).

Stimulus

The stimulus for the Galilean Moons cluster is shown in Figure 17.

Figure 17. Stimulus: Galilean Moons

Four of Jupiter's closest moons can be seen orbiting the planet by using a low-powered telescope. A ruler on the lens of the telescope is used to take measurements. The animation shows the movements of the moons and Jupiter over the course of several days. Click on the small gray arrow at the bottom left of the picture to begin the animation.



The table shows data on each of the moons.

Data on Galilean Moons

	Diameter (km)	Mean Distance from Jupiter (km)	Orbital Period (days)
Callisto	4,800	2,000,000	16.7
Europa	3,318	700,000	3.5
Ganymede	5,262	1,000,000	7.2
Io	3,630	400,000	1.8

Details by Item

Item 1

Item 1 of the Galilean Moons cluster is shown in Figure 18.

Figure 18. Item 1: Galilean Moons

Use the measuring tool on the animation to determine each moon's maximum distance from Jupiter.

Complete the table by entering the measurements to the closest 0.25 mark.

	Maximum Distance from Jupiter in Animation
M1	
M2	
M3	
M4	

SCORES

This item was relatively easy for students; six students earned 4 score points (full credit), and one other student earned 3 score points. However, four students earned no credit (including one student who skipped over the item without attempting to answer it).

Eight of the 12 students seemed comfortable manipulating the simulation and re-watched, with appropriate pauses, to figure out each moon's distances from Jupiter. Some also re-watched the simulation while responding to Item 2.

One student neglected to watch the simulation at all.

COMPREHENSION

Although, the introduction to the stimulus states that "A ruler on the lens of the telescope is used to take measurements," five students did not understand the measuring tool, or the units used on the tool.

> One of these students used the mean distance from Jupiter in kilometers from the *Data on Galilean Moons* table for her responses to the item. The student said that the instructions suggested using a measuring tool, but she did not see a measuring tool.
>
> Another student said, "I thought the numbers [going across the lens on the animation] were extremely confusing. I think that if they're trying to take it to orbital days, then they have to make the length longer, but if it takes 16.7 days—well that's orbit. I don't know, it's just super confusing. They should say that the numbers represent the length of time or the number of days."

At least two students were confused by the instructions "to the closest 0.25 mark."

REASONING

The seven students who earned three or 4 score points all showed evidence in the think-aloud of using the animation in the manner intended to formulate their response.

> For example, one student said that she was going to follow one moon at a time "because I can't follow all of them at the same time." As she watched the animation a second time, she noted where each of the moons was, narrating aloud, "M2 is around the 1.5 mark. M4 is around the 2.5 mark." She then paused the video, studied the text of Item 1, and began entering the data. When she reached the response field for M3, she said, "I'll just leave it at 7, because it went a little past 7 but not too far."

Item 2

Item 2 of the Galilean Moons cluster is shown in Figure 19.

Figure 19. Item 2: Galilean Moons

Select the boxes to identify each moon by name.

	Callisto	Eurpoa	Ganymede	Io
M1	☐	☐	☐	☐
M2	☐	☐	☐	☐
M3	☐	☐	☐	☐
M4	☐	☐	☐	☐

Scores

This item was also relatively easy for students; seven students received full credit (4 score points), and only one student received no credit.

Comprehension

No features of this item appeared to confuse students.

Reasoning

Nearly all the students reasoned their way to a solution using the stimulus materials as intended.

> For example, one student stated she was going to look for the mean distance from Jupiter [on the *Data on Galilean Moons* table] and use what she got from the previous question—the maximum distance for each moon. The student selected M3 for Callisto "because it is the farthest away and has the largest mean distance." She noted that Europa has the third "biggest" mean and, looking for the third largest maximum distance, deduced that M4 must be Europa. Seeing that Ganymede has the second largest mean distance, the student selected M1. The last moon left (Io) was identified by default as M2.

Item 3

Item 3 of the Galilean Moons cluster is shown in Figure 20.

Figure 20. Item 3: Galilean Moons

Compare the measurements you took to the distances in the Data on Galilean Moons table. Then, select the statement that is true.

Ⓐ The measurements you took are proportional to the data in the table.

Ⓑ The measurements you took are not proportional to the data in the table because the table is wrong.

Ⓒ There is not enough information to tell whether the measurements you took are proportional to the data in the table.

Ⓓ The data you measured is not proportional to the data in the table because your measurement instrument is imprecise at that distance.

SCORES

This item was much more challenging than the other items in the cluster, and only three students selected the correct response that the data the student measured are not proportional to the data in the table due to the differences in measurement accuracy.

The nine students who did not earn credit for this item were fairly evenly distributed across the distractors (four students chose C, three chose A, and two chose B), suggesting that they really were at a loss to understand how to explain the differences between their measurements and the data in the table.

COMPREHENSION

Two students said that they did not know the meaning of "proportional," and, based on the item responses, it's likely that a number of others did not fully understand the concept of proportional.

Although not mentioned, students may also not have understood what it meant that "your measurement instrument is imprecise."

REASONING

Even students who selected the right answer, may not have done so with full comprehension.

> For example, one student read through all the answers, then started eliminating answers. First, she eliminated A and B, then decided the answer was D because the ruler measured the distance in the animation, but the table gave the distances in kilometers.

3.3.2 Cluster 3: Hippos

Performance Summary

The median time to complete the Hippos cluster was 10 minutes. Table 23 and Table 24 indicate the number of students attaining cluster total scores and item scores within the specified ranges, respectively.

Table 23. Number of Students Attaining Cluster Total Scores in Specified Range: Hippos

Score 10–7	Score 6–4	Score 3–1	Score 0
2	5	3	0

Note. Maximum score = 10; *n* = 10; two students ran out of time before completing this cluster.

Table 24. Number of Students Attaining Item Scores in the Specified Range, by Item: Hippos

	Maximum Item Score	Score 4–3	Score 2–1	Score 0
Item 1	4	1	9	0
Item 5	3	1	4	5

	Maximum Item Score	Score 1	Score 0
Item 2	1	5	5
Item 3	1	7	3
Item 4	1	3	7

Note. n = 10; two students ran out of time before completing this cluster.

Task Demands

The following are task demands of the Hippos cluster:

- Articulate, describe, illustrate, or select the relationships or interactions to be explained. This may entail sorting relevant from irrelevant information or features (Item 1).
- Express or complete a causal chain common or distinct across organisms or environments. This may include indicating directions of causality in an incomplete model such as a flow chart or diagram or completing cause and effect chains (Item 2).
- Express or complete a causal chain common or distinct across organisms or environments. This may include indicating directions of causality in an incomplete model such as a flow chart or diagram or completing cause and effect chains (Item 3).

- Articulate, describe, illustrate, or select the relationships or interactions to be explained. This may entail sorting relevant from irrelevant information or features (Item 4).
- Use an explanation to predict interactions among different organisms or in different environments (Item 5).

Stimulus

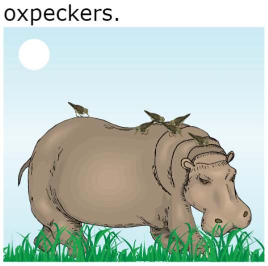
The stimulus for the Hippos cluster is shown in Figure 21.

Figure 21. Stimulus: Hippos

In Africa, a variety of organisms coexist with others in distinct ecosystems. For example, hippopotamuses spend time in both aquatic and savannah ecosystems.

When found in aquatic environments, hippopotamuses are often surrounded by carp.

When found in a savannah environment, hippopotamuses are often surrounded by birds called oxpeckers.

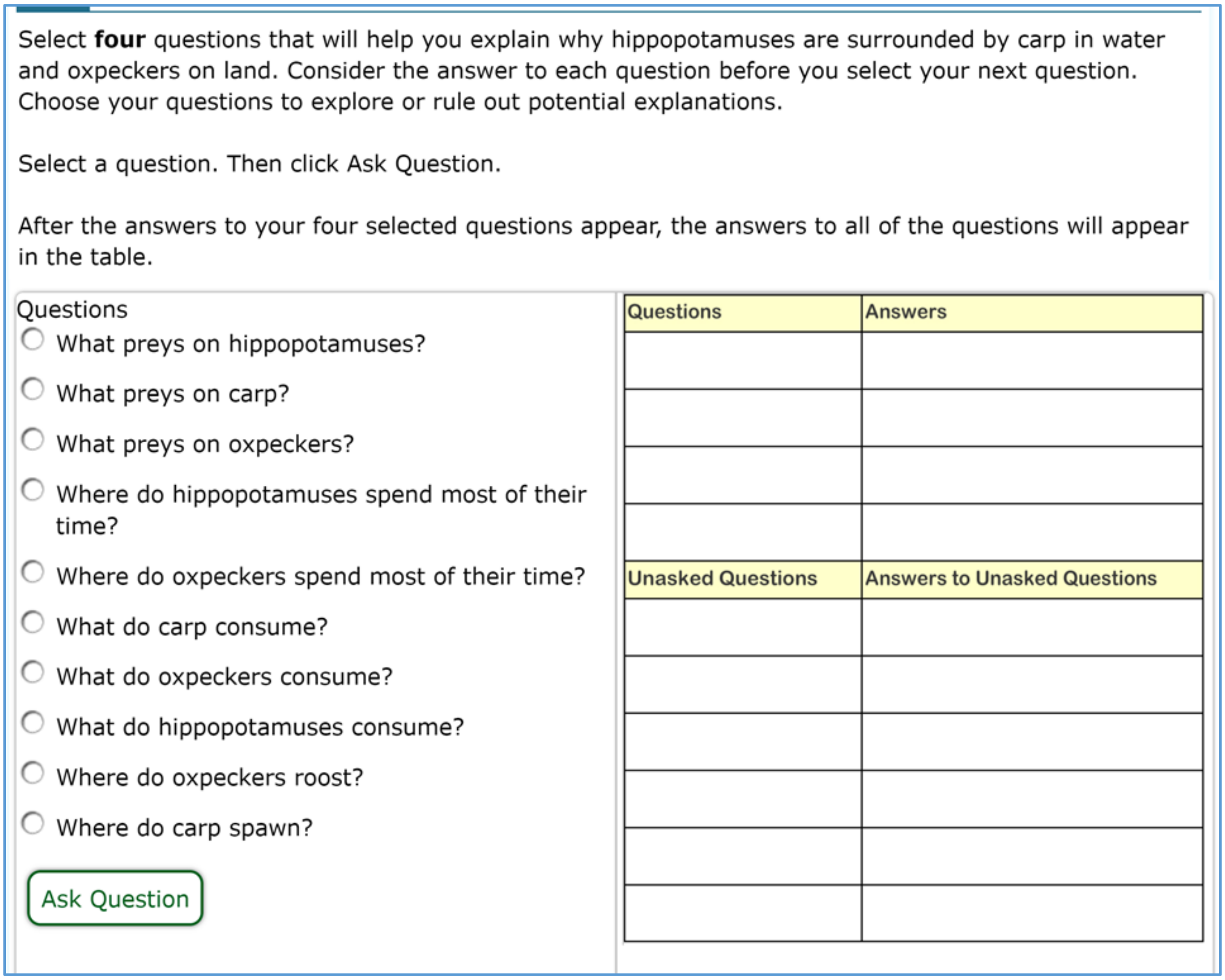


Details by Item

Item 1

Item 1 of the Hippos cluster is shown in Figure 22.

Figure 22. Item 1: Hippos



SCORES

Every student earned some credit on this item:

- One student earned 4 points (full credit).
- Three students earned 3 points.
- Six students earned 2 points.
- One student earned 1 point.

COMPREHENSION

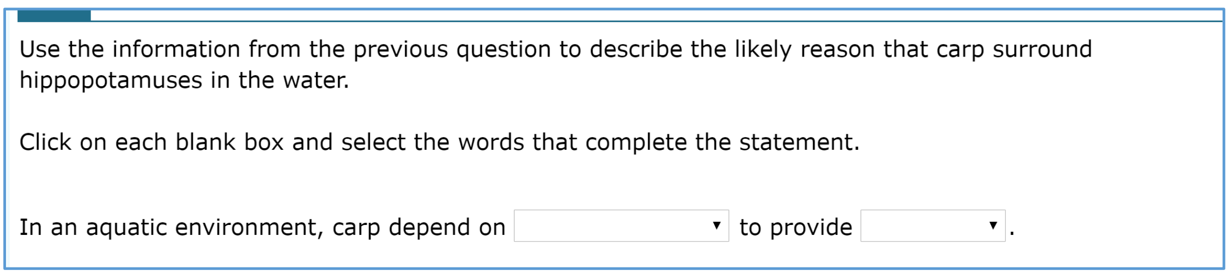
As evidenced from their reasoning in the think-alouds, students understood that they were to choose questions they thought would be helpful to explain the relationships between hippos and oxpeckers or carp, although, as can be seen from the score distribution, they did not necessarily know what those questions would be. Two students, however, commented on the fact that being asked to choose questions seemed like a waste of time in light of the fact that answers eventually were populated for all the questions.

Three students did not initially understand that they had to click "Ask Question" and could only ask one question at a time; one student initially thought that she had to type the text of the question rather than select from the list.

Item 2

Item 2 of the Hippos cluster is shown in Figure 23.

Figure 23. Item 2: Hippos



SCORES

Half of the students (five) received credit for this item.

COMPREHENSION

Students found this item easy to comprehend, and they had sufficient knowledge of transactional relationships among animals to understand the concept behind the item.

Score variance on this item (and the next) came from the "to provide" response; students found it obvious that the response for the first drop-down box should be Hippopotamuses.

REASONING

Most students reasoned appropriately from the information in Item 1 to determine their response.

> For example, one student said, "In an aquatic environment, carp depend on . . . so why would a carp depend on the hippopotamus? [Referring back to question 1:] So what preys on hippos? I don't need that. Where do they spend their time? I don't need that. Where do oxpeckers spend most of their time? On the bodies of host mammals. What do hippos consume? Grass and plants. Where do oxpeckers roost? On the bodies of host mammals. Oh, so I believe that in the aquatic environment, carp depend on hippos to provide . . . food . . . Because they eat fleas, dead skin, parasites, and mucous."

Those who did not respond correctly simply made wrong inferences from the data—some of which were wrong but plausible.

> For example, one student explained why he selected protection by saying, "hippopotamuses are a much bigger animal than the fish and could provide protection from the crocodile." The student noted that, in Item 1, one of the answers indicated that crocodiles, snakes and larger fish prey on carp.

Item 3

Item 3 of the Hippos cluster is shown in Figure 24.

Figure 24. Item 3: Hippos

Use the information from the previous question to describe the **most likely** reason that oxpeckers surround hippopotamuses on the land.

Click on each blank box and select the words that complete the statement.

In the savannah environment, oxpeckers depend on [▾] to provide [▾].

Scores

Seven students received credit for this item.

Comprehension

This item is very similar to Item 2, and the same observations about comprehension apply.

Reasoning

This item is very similar to Item 2, and the same observations about reasoning apply.

Item 4

Item 4 of the Hippos cluster is shown in Figure 25.

Figure 25. Item 4: Hippos

Select the boxes to identify which organisms are paired with the hippopotamus in the described relationships.

	Oxpecker	Carp	Neither
Predatory relationship	☐	☐	☐
Competitive relationship	☐	☐	☐
Mutually beneficial relationship	☐	☐	☐

SCORES

Three students earned credit on this item, which required that all three answers about organisms in relationships with hippos be correct. The fewest students (two) correctly identified the answer for *Competitive relationship*.

COMPREHENSION

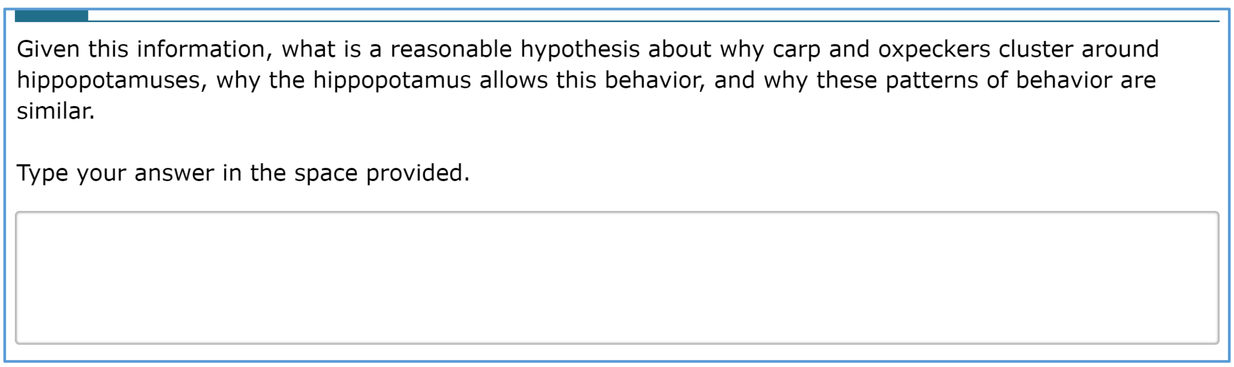
Although students generally understood the concept of transactional relationship among animals, some lacked prior knowledge of the terms used in the item.

> For example, one student said that "mutually beneficial" was the only relationship mentioned in the sample lesson. He did not know if the predatory and competitive relationships were "interchangeable or how it worked."

Item 5

Item 5 of the Hippos cluster is shown in Figure 26.

Figure 26. Item 5: Hippos



Given this information, what is a reasonable hypothesis about why carp and oxpeckers cluster around hippopotamuses, why the hippopotamus allows this behavior, and why these patterns of behavior are similar.

Type your answer in the space provided.

SCORES

One student earned full credit (3 score points) by providing correct hypotheses for each of the three questions posed in the item stem.

Four other students provided a correct hypothesis for at least one of the questions.

COMPREHENSION

There were no comprehension issues with this item.

REASONING

Some students failed to address the task of formulating hypotheses altogether. Others made appropriate use of the information gathered from the previous items in formulating their responses, but, given that their understanding of the previous items was not necessarily correct, these misunderstandings could carry over into this item.

3.3.3 Cluster 3: Morning Fog

Performance Summary

The median time to complete the Morning Fog cluster was 12 minutes. Table 25 and Table 26 indicate the number of students attaining cluster total scores and item scores within the specified ranges, respectively.

Table 25. Number of Students Attaining Cluster Total Scores in Specified Range: Morning Fog

Score 9–7	Score 6–4	Score 3–1	Score 0
2	3	7	0

Note. Maximum score = 9; *n* = 12.

Table 26. Number of Students Attaining Item Scores in Specified Range, by Item: Morning Fog

	Maximum Item Score	Score 7–6	Score 5–3	Score 2–1	Score 0
Item 1 (Parts A–C)	7	0	10	2	0

	Maximum Item Score	Score 2	Score 1	Score 0
Item 1 (Part D)	2	3	0	9

Note. n = 12.

Task Demands

The following are task demands of the Morning Fog cluster:

- Select or identify from a collection of potential model components including distractors, the components needed to model the model of evaporation, condensation, transpiration, precipitation, or other behaviors of water molecules during the water cycle.
- Assemble or complete, from a collection of potential model components, an illustration or flow chart that represents the phenomenon. This does not include labeling an existing diagram.
- Given models or diagrams of the phenomenon, identify the parts of the model and how they change in each scenario OR identify the properties of the model that cause the change.

Stimulus

The stimulus for the Morning Fog cluster is shown in Figure 27.

Figure 27. Stimulus: Morning Fog

Morning Fog in a Valley

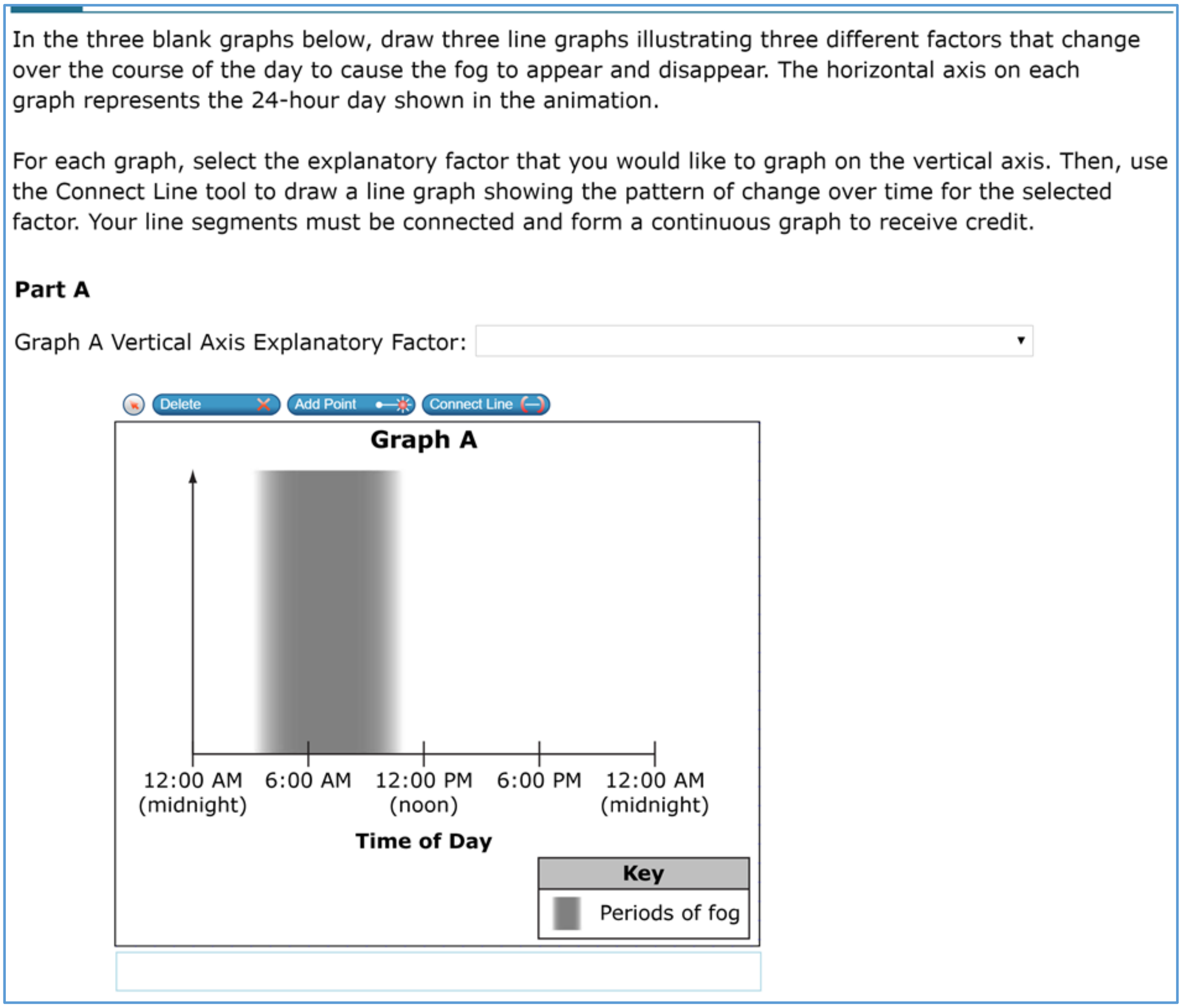
Fog appears and disappears over the course of the morning in the Willamette Valley in Oregon. The animation shows the appearance and disappearance of fog in the valley during a 24-hour day. The sun rises at 6 AM and later sets at 6 PM.

Details by Item

Item 1

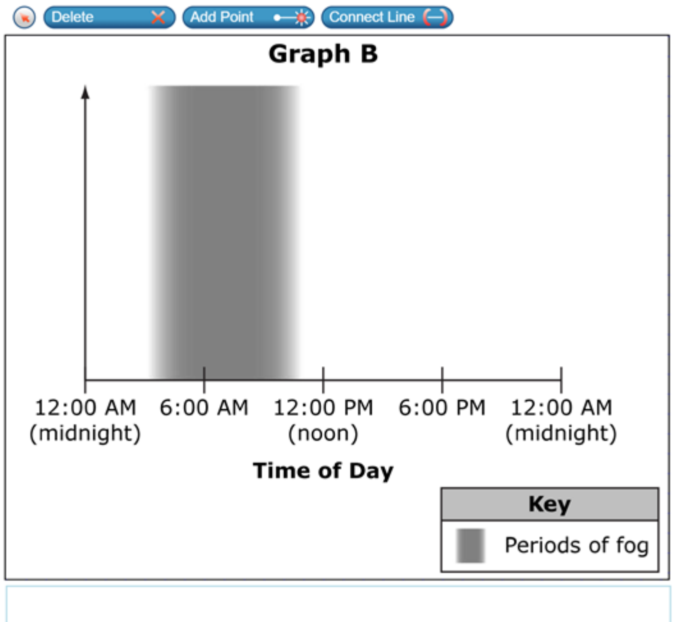
Item 1 of the Morning Fog cluster is shown in Figure 28.

Figure 28. Item 1: Morning Fog



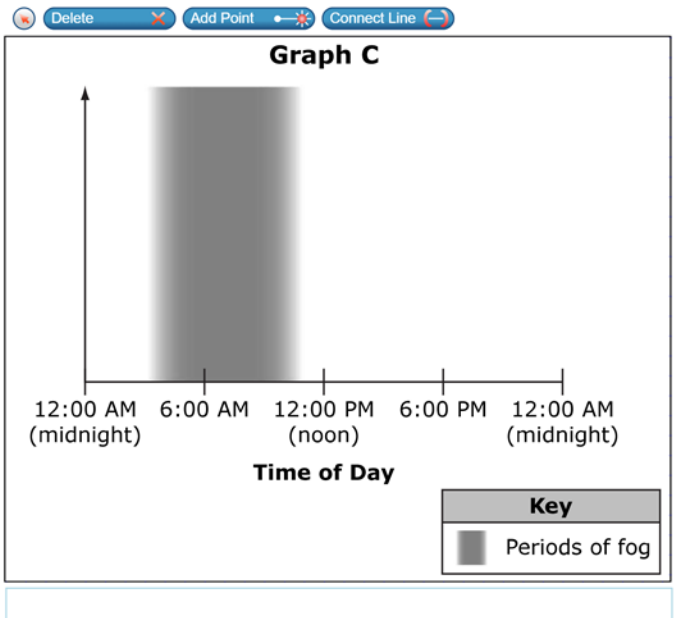
Part B

Graph B Vertical Axis Explanatory Factor: [dropdown]



Part C

Graph C Vertical Axis Explanatory Factor: [dropdown]



Part D

The process described in [dropdown] causes the process described in [dropdown], which causes the process described in [dropdown].

Item 1 (Parts A–C)

SCORES

Parts A–C were scored as a unit.

Students could earn up to 6 points for correctly drawing three-line graphs showing how weather factors affecting fog formation changed over the course of the day; they could earn up to 3 points for correctly identifying the explanatory factor associated with each of the processes they chose to graph.

Half of the students (six) earned some credit for their graphs, but none earned full credit.

- Six earned points for graphing a decrease in the evening in one or more of the following: sunlight intensity, temperature, and/or proportion of water in the air
- Six earned points for graphing sunlight intensity, showing both an increase in the morning and a decrease in the evening.

No one earned points for graphing either the proportion of water in the air declining as the fog forms and increasing as the fog dissipates, or the temperature decreasing when the fog begins to form and rising when the fog dissipates.

Four students did not earn any credits for their graphs, and their graphs did not resemble the correct answers: they included horizontal lines, a single line that ascended, and dots with no connecting line.

All but two of the students earned at least two out of the three possible score points for the explanatory factors. The numbers of students earning points for correctly identifying each explanatory factor were as follows:

- Sunlight intensity (nine students)
- Air temperature (eight students)
- Proportion of water in the air in gas form (nine students)

COMPREHENSION

Eight students were confused about how to draw the line graphs, including four who did not understand that they had to define the value of the y-axis. The following are examples of think-alouds from students who were confused by the graphs:

- “I have no idea. I don’t understand this graph. It’s confusing. Since there’s nothing on the left, the vertical. (referring to the y-axis). The three factors that can change, I have no idea what they mean by that. I feel like they’re not giving enough information for me to understand. I’m so confused. The three different factors are what—the nighttime? What’s the difference between the graphs? Wouldn’t they all be the same? Oh, three different factors.” (The student apparently didn’t see the explanatory factor drop-down menu until this point.)

- The student re-read the part of the question that discusses "showing the pattern of change over time for the selected factor" and commented, "yeah, that really doesn't make sense, how they want me to connect the line. If I saw this on a test, I would just freak out because I wouldn't know how I was supposed to draw a line graph to represent this."
- "How do you represent how much fog? I'm guessing"—the student clicked to create some points—"I'm guessing it'd be something like that." The student clicked around some more and then connected the points. "I guess that's what I'm gonna say, because this really doesn't make sense how they want you to draw a graph. If anything, they should have increments and a chart of how high the fog rises or how much of whatever is in the air."

Six students were initially unclear about how to use the pull-down menu of explanatory factors, but mostly figured out how to use them.

Two students had a somewhat better understanding of Parts A–C after they read Part D and went back and changed some of their answers in Parts A–C.

> For example, after reading Part D, one student realized that each graph was meant to represent a different factor. When asked, the student said that he misunderstood the question and picked the same factor for all three graphs at first because he didn't know what was meant by the term "explanatory factor," and thought the question was just asking about the fog.

REASONING

Half of the students (six) re-watched the animation while drawing the line graphs.

An example of correct reasoning from the animation comes from the student who earned the most score points on parts A–C (7 points). She indicated that she chose Proportion of Water in the Air for her first graph because it was "the one that related to the fog the most." When asked to explain more about her graph, the student said she looked at the animation "to see the intensity of the fog and when it decreased" and that's why she made the graph increasing then decreasing. "First increasing from 3 to 6 [A.M.], then decreasing from 6 to 8."

<u>Item 1 (Part D)</u>

SCORES

Only three students earned the two possible core points by correctly responding that variations in sunlight intensity affect air temperature, which, in turn, affects the proportion of water in the air in gas form (water cycle).

COMPREHENSION

Since most students were confused by Parts A–C, they also had trouble understanding what they were being asking to do in Part D.

3.3.4 Cluster 4: Texas Weather

Performance Summary

The median time to complete the Texas Weather cluster was 14 minutes. Table 27 and Table 28 indicate the number of students attaining cluster total scores and items scores within the specified ranges, respectively.

Table 27. Number of Students Attaining Cluster Total Scores in Specified Range: Texas Weather

Score 11–7	Score 6–4	Score 3–1	Score 0
0	4	8	0

Note. Maximum score = 11; *n* = 12.

Table 28. Number of Students Attaining Item Scores in Specified Range, by Item: Texas Weather

	Maximum Item Score	Score 8–7	Score 6–4	Score 3–1	Score 0
Item 1 (Part A)	8	0	2	8	2

	Maximum Item Score	Score 1	Score 0
Item 1 (Part B)	1	1	11
Item 2	1	4	6
Item 3	1	6	3

Note. n= 12 for Item 1, Parts A and B; 11 for Item 2, and 10 for Item 3. One student did not scroll down to Items 2 and 3, and one student gave up and refused to attempt Item 3.

Task Demands

The following are task demands of the Texas Weather cluster:

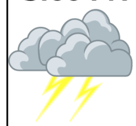
- Describe, illustrate, or select tools, locations, and/or methods to use in investigations of phenomena related to interactions of air masses. This should show how or where measurements will be taken (Item 1).
- Identify, select, or describe the relevance of particular data or sources relevant to the process of weather forecasting (Item 1).
- Predict the effects of given changes in the air masses' interactions on subsequent weather (Item 2).
- Predict the effects of given changes in the air masses' interactions on subsequent weather (Item 3).

Stimulus

The stimulus for the Texas Weather cluster is shown in Figure 29.

Figure 29. Stimulus: Texas Weather

The weather in Austin turned cold and wet around 3:00 p.m. yesterday. Following is the hour-by-hour weather report for Austin.

	Noon	1:00 PM	2:00 PM	3:00 PM	4:00 PM	5:00 PM
						
Temperature	80° F	75° F	70° F	68° F	66° F	65° F
Chance of rain	0%	30%	50%	95%	100%	100%
Humidity	80%	85%	88%	92%	95%	96%
Wind	SE 9 MPH	SE 10 MPH	SE 9 MPH	NW 12 MPH	NW 13 MPH	NW 12 MPH
Pressure	32.0 inHG	30.3 inHG	29.9 inHG	29.0 inHG	28.7 inHG	28.5 inHG

As you work through the following questions, you will gather the information needed to explain the cause of this weather pattern.

Details by Item

Item 1

Item 1 of the Texas Weather cluster is shown in Figure 30.

Figure 30. Item 1: Texas Weather

Part A

The following question has two parts. First, answer part A. Then, answer part B.

Use the simulator to take measurements that will help you determine what caused Austin's afternoon weather.

You will be scored on your selections, so be sure to:

- specify what you are looking for,
- use the appropriate tools to look for them,
- keep taking measurements until you know what caused the weather, and
- stop taking measurements when you have all the information you need.

You may take a maximum of 8 measurements.

Checking for a(n) Air Mass

Location 1

Time of day 3pm

Tool 1 Thermometer

Tool 2 Barometer

Take Measurement

Measurement Number	Location	Checking For	Time of Day	Temperature	Wind Speed	Wind Direction	Pressure

Part B

From the measurements that you have taken, indicate up to two measurements (by "Measurement Number" from the result table in the simulation) that provide sufficient evidence for the claim in the first column. Be sure to select "None" if the measurements do not provide sufficient evidence of a claim.

	1	2	3	4	5	6	7	8	None
A low pressure air mass moved west towards Austin.	☐	☐	☐	☐	☐	☐	☐	☐	☐
A high pressure front moved south towards Austin.	☐	☐	☐	☐	☐	☐	☐	☐	☐
A cold front moved north towards Austin.	☐	☐	☐	☐	☐	☐	☐	☐	☐
Precipitation moved into Austin from the east.	☐	☐	☐	☐	☐	☐	☐	☐	☐

Item 1 (Part A)

Scores

Part A was extremely difficult for students, and the randomness of earned points across students suggests that none of the students really understood what they were supposed to do with the simulator, either because they didn't have the requisite content knowledge or they were confused by the manner in which the simulator was presented.

Four of the points in the scoring rubric for Part A involve the parameters that the student chooses for trials on the simulator or matching the right tools with the right parameters, but many students failed to change the parameter on successive trials and simply focused on manipulating the tools. Four students used air mass (the default) for all of their measurements, and two students used primarily air mass. Consequently, score points based on choice of parameter or match between parameter and tools may not be meaningful. That said,

- nine students earned 1 score point for selecting air mass as the parameter on at least one trial;
- no students earned a score point for matching the correct tools with air mass;
- no students earned a score point for selecting movement as the parameter; and
- two students earned a score point for matching the correct tools with movement on at least one trial.

The four remaining points for Part A were awarded for measuring the correct factor at the proper locations and/or time and for doing so using the correct tools.

- Three students earned a point for at least one trial checking for movement measured at locations 3, 4, or 5.
- A different student earned a point for at least one trial checking for air mass measured at 1 p.m. at locations 3, 4, or 5.

The criterion statements in this section of the rubric were inconsistent. The criterion on which three students earned a point was the most permissive in that it specified a location, but not a time.

Comprehension

Seven students did not initially understand what actions they were supposed to take to run trials on the simulator. Seven other students were unfamiliar with some of the measuring tools and did not know what they measured. Another student took only one measurement because he did not understand how to take more measurements.

The instructions to "determine what caused Austin's afternoon weather" were too open ended for these students.

- At least three students noted that the answer choices in Part B would have given them an idea of how to tackle the problem if they had read Part B before working with the simulator.

- Two students earned the most credits on Part A (4 score points) by (1) checking for air mass and movement, (2) choosing wind vane and anemometer when checking for movement, and (3) conducting one trial for air mass measured at 1 p.m. at locations 3, 4, and 5. One of these students said she was confused and overwhelmed when probed about this item.
    - "There was no way I could read this and understand it, I'll just look back and forth between [the chart and the table]." The student explained, "I've never been good with weather – it doesn't make sense to me how everything works . . . I didn't understand the table – like how it correlated with what I was putting in [Part A]. I was overwhelmed with eight measurements because it said, 'Do Part A and then Part B,' so I was thinking okay, I should do Part A and then Part B. But then after I did Part B, I realized that I should have looked at Part B first so I would know what eight measurements to take! I didn't know the difference in what would show up on the table if I chose air mass, or movement, or precipitation. I just didn't understand what difference it would make in each choice I had."

REASONING

The other student who earned 4 score points on the item had a somewhat better understanding of how to use the simulator to find out what caused Austin's afternoon weather.

> In her think-aloud, the student said that she was going to take measurements first at Location 3 because it's most central. She chose 3 p.m. because that's when the weather turned cold and wet in Austin. She then changed the measurement to Location 4 because "it's closest to Austin and what the chart pertains to." Said she would leave the time as 3 p.m. as that's when it was cold and wet. She said she would use the anemometer and the thermometer. She clicked *Take Measurement*. She said she would check for precipitation but didn't see any tools that pertained. She then chose movement at Location 3, using a wind vane and an anemometer, to see if the wind was going in that direction.

<u>Item 1 (Part B)</u>

SCORES

Only one student got credit for Part B, and this may have been by chance, given that the student only earned one of the eight possible points on Part A.

COMPREHENSION

At least three students did not realize that the numbers 1 through 8 on Part B were the eight measurements they were allowed to take in Part A, and that they were to pick measurements that showed evidence for the claim in column 1.

REASONING

Given their performance on Part A, students had little to work with in Part B, even if they understood what they were supposed to do.

> For example, one student said that she had to make her best guess in Part B because "none of my measurements in Part A told me anything because I took all the wrong measurements in Part A. Part B was truly kind of stressful for me."

Item 2

Item 2 of the Texas Weather cluster is shown in Figure 31.

Figure 31. Item 2: Texas Weather

Suppose that it was hot and humid in San Antonio at 3:00 p.m. What does the pattern of weather suggest for precipitation in San Antonio in the evening?

- (A) The pattern is not likely to affect precipitation in San Antonio in the evening.
- (B) The pattern suggests that the chance of rain in San Antonio will stay about the same as it was at 3:00 p.m.
- (C) The pattern suggests that the chance of rain will increase.
- (D) The pattern suggests that the chance of rain will decrease.

Scores

Four of the 10 students who attempted this item earned credit.

Comprehension

Given performance on Item 1, it is unlikely that these students' scores actually reflected mastery of the content being assessed by the item.

Some students understood "pattern of weather" as referring to the hour-by-hour weather report shown in the stimulus, and it's not clear that any of the students realized that the question pertained to a different location than the weather report (or Item 1).

> For example, one student referred to the weather report table and said that the table indicates that the chance of rain will likely increase so he couldn't select decrease (pointing at both option A and option D). The student noted that option B suggests no change, but the table shows a very clear change in the chance of rain, therefore B could not be the answer. The student referred to the table again and said that the chance of rain was increasing, so C was the only possible answer that works.

Item 3

Item 3 of the Texas Weather cluster is shown in Figure 32.

Figure 32. Item 3: Texas Weather

Suppose that it was hot and humid in San Antonio at 3:00 p.m. What does the pattern of weather suggest for the temperature in San Antonio in the evening?

Ⓐ The pattern is not likely to affect temperature in San Antonio in the evening.

Ⓑ The pattern suggests that temperature in San Antonio will stay about the same as it was at 3:00 p.m.

Ⓒ The pattern suggests that the temperature will increase.

Ⓓ The pattern suggests that the temperature will decrease.

Scores

Six of the nine students who attempted this item earned credit.

Comprehension

As with the other items in this cluster, students had, at best, a faulty understanding of this item. Consequently, as with Item 2, a correct response did not indicate mastery of the content being assessed.

> For example, one student said that, as soon as she read "temperature," she went to the weather report table, looked at the temperature at 3 p.m., and saw that the temperature was decreasing over time. The student then went back to the question and read through the options and noted that answer A was about no effect, that B was about staying the same, and C was about the temperature increasing. Since the temperature is decreasing, the student decided that answer D was the only one that matched the data.

3.4 DETAILED DISCUSSION BY CLUSTER: HIGH SCHOOL

3.4.1 Cluster 1: Blood Sugar Regulation

Performance Summary

The median time to complete the Blood Sugar Regulation cluster was 19 minutes. Table 29 and Table 30 indicate the number of students attaining cluster total scores and item scores within the specified ranges, respectively.

Table 29. Number of Students Attaining Cluster Total Scores in Specified Range: Blood Sugar Regulation

Score 7–6	Score 5–3	Score 2–1	Score 0
0	9	3	1

Note. Maximum score = 7; *n* = 13; two students ran out of time before completing this cluster.

Table 30. Number of Students Attaining Item Scores in Specified Range, by Item: Blood Sugar Regulation

	Maximum Item Score	Score 3	Score 2–1	Score 0
Item 1	3	8	4	1
Item 2	3	0	3	11

	Maximum Item Score	Score 2	Score 1	Score 0
Item 3	2	3	7	3

Note. *n* = 13; two students ran out of time before completing this cluster.

Task Demands

The following are task demands of the Blood Sugar Regulation cluster:

- Identify the outcome data that should be collected in an investigation to provide evidence that feedback mechanisms maintain homeostasis. This could include measurements and/or identifications of changes in the external environment, the response of the living system, stabilization/destabilization of the system's internal conditions, and/or the amount of systems for which data is collected.
- Make and/or record observations about the external factors affecting systems interacting to maintain homeostasis, responses of living systems to external conditions, and/or stabilization/destabilization of the system's internal conditions.

- Identify or describe the relationships, interactions, and/or processes that contribute to and/or participate in the feedback mechanisms maintaining homeostasis that lead to the observed data.
- Using the collected data, express or complete a causal chain explaining how the components of (a) mechanism(s) interact in response to a disturbance in equilibrium in order to maintain homeostasis. This may include indicating directions of causality in an incomplete model such as a flow chart or diagram or completing cause and effect chains.
- Evaluate the sufficiency and limitations of data collected to explain the cause and effect mechanism(s) maintaining homeostasis.

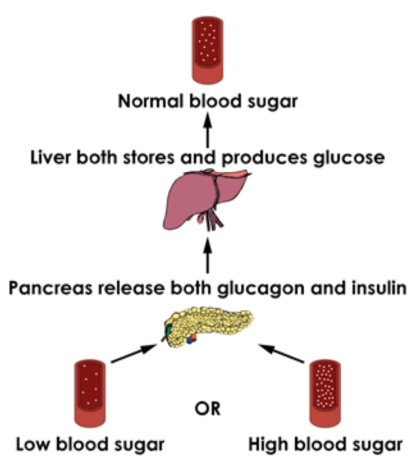
Stimulus

The stimulus for the Blood Sugar Regulation cluster is shown in Figure 33.

Figure 33. Stimulus: Blood Sugar Regulation

A hungry person eats a meal. Soon after the meal is completed, the person's blood sugar is elevated. After a while, the blood sugar levels return to their pre-meal levels.

Hunger is one of the body's symptoms of abnormal blood glucose levels, or blood sugar. Hunger alerts the body to eat, which almost immediately increases blood sugar. Both the pancreas and liver work together to maintain blood sugar concentrations in the range of 80-120 miligrams per deciliter (mg/dL). The pancreas helps regulate blood sugar by producing two types of hormones: glucagon and insulin. The normal range for blood glucagon levels is 60-200 picograms per mililiter (pg/mL) and the normal range for blood insulin levels is 65-200 picomole per liter (pmol/L). The liver both converts glucagon into glucose and stores glucose. The flowchart shows how the pancreas and liver participate in feedback mechanisms to help regulate blood sugar.



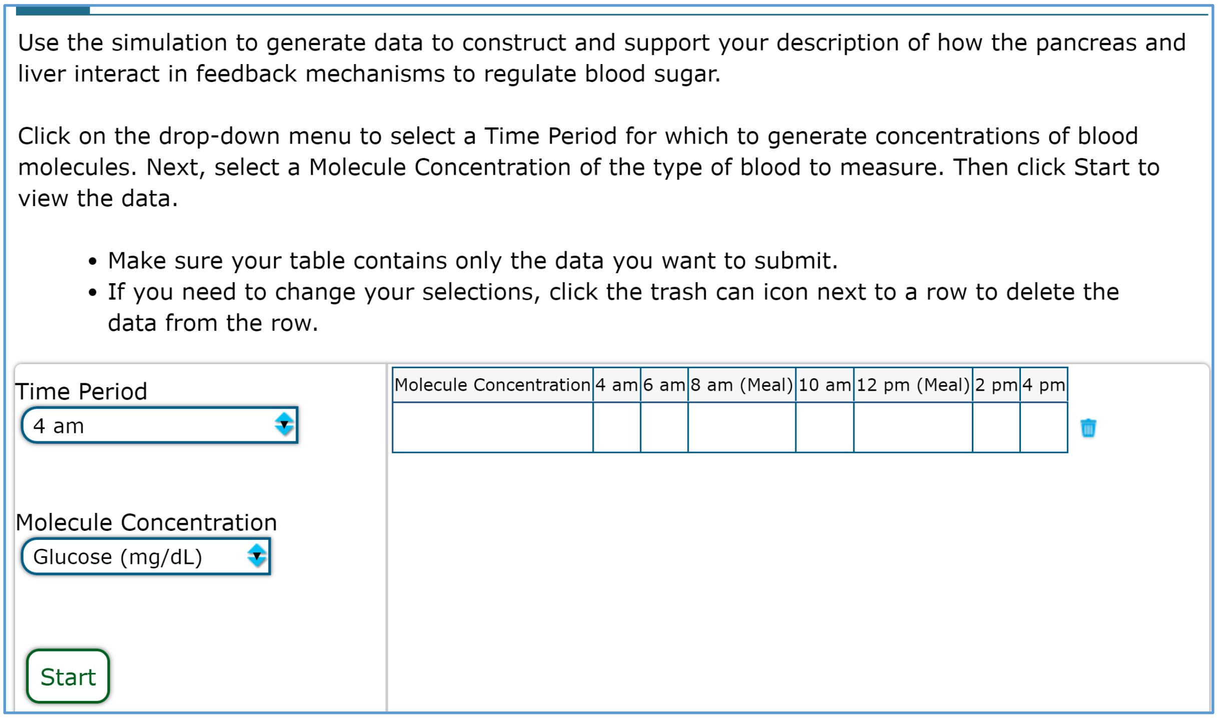
In the questions that follow, investigate and describe how the molecules produced and stored by the pancreas and liver interact in feedback mechanisms to regulate blood sugar.

Details by Item

Item 1

Item 1 of the Blood Sugar Regulation cluster is shown in Figure 34.

Figure 34. Item 1: Blood Sugar Regulation



Scores

Student scores on this item are as follows:

- Eight students earned 3 score points (full credit).
- Three students earned 2 score points.
- Two students earned 1 score point.

Comprehension

Seven students expressed some confusion in figuring out how to generate data in the simulation. For example, one student was confused by the layout of the item and by the term "simulation" because she was not sure whether she should test all the options or provide her own answer. At this point she skipped ahead to look at the next items to see if they would provide any clues as to how she should proceed on Item 1 but did not find that helpful. She was very unsure what to do next and seemed overwhelmed by the options. After some flipping back and forth, she decided to measure all three values for each of the times offered.

At least three students went back to Item 1 and re-generated the data in the simulation once they knew that they had to create three graphs in Item 2.

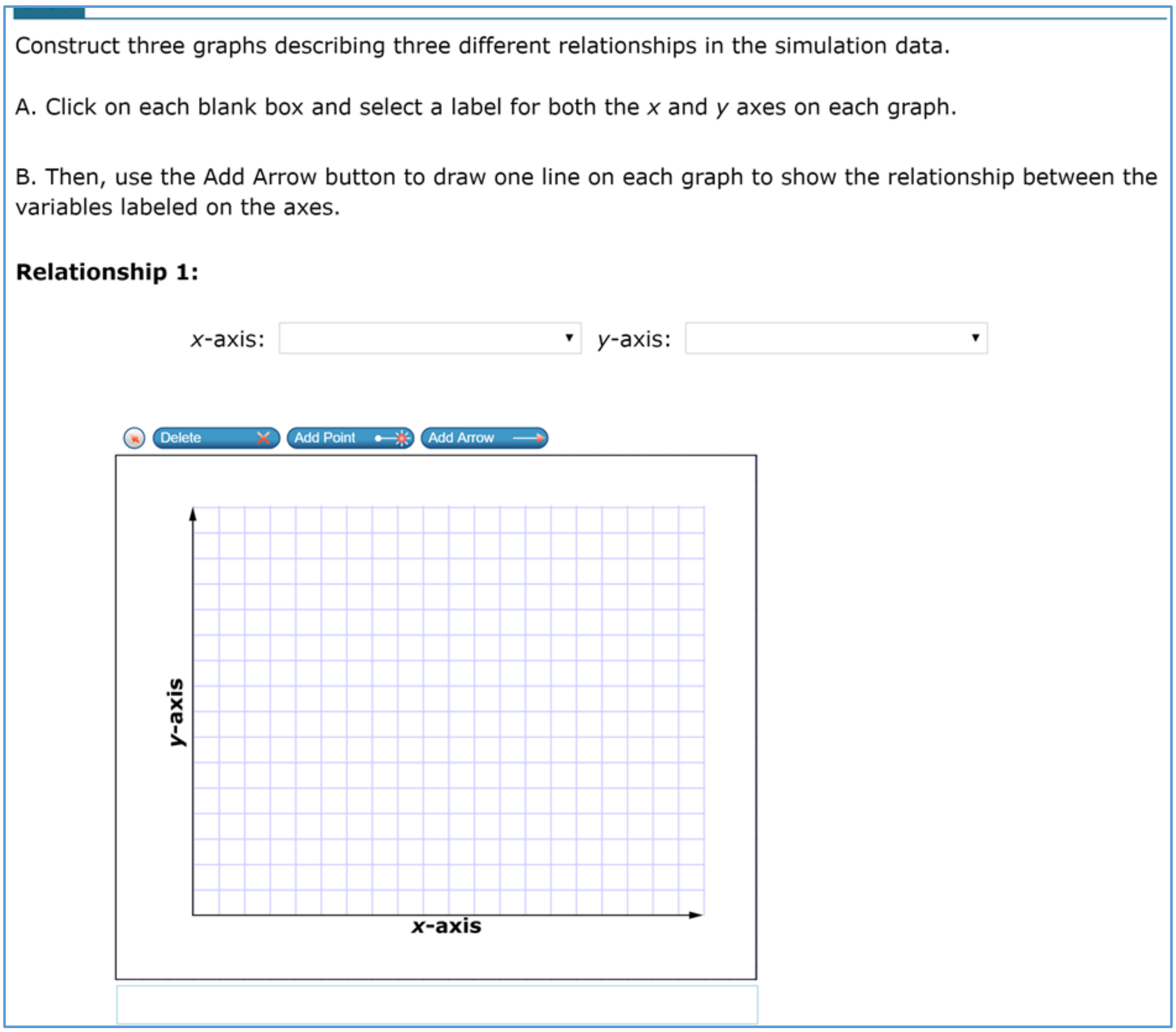
Reasoning

Students used the simulations as a learning experience. For example, when asked how he decided how many simulations to do, one student said, “Well, I knew that there was three different substances (glucose, glucagon, and insulin). I wasn’t really sure how it worked, and then once I did it, I was like ‘OK well that’s when you have a meal,’ so I knew from the reading that’s when your blood sugar spikes.”

Item 2

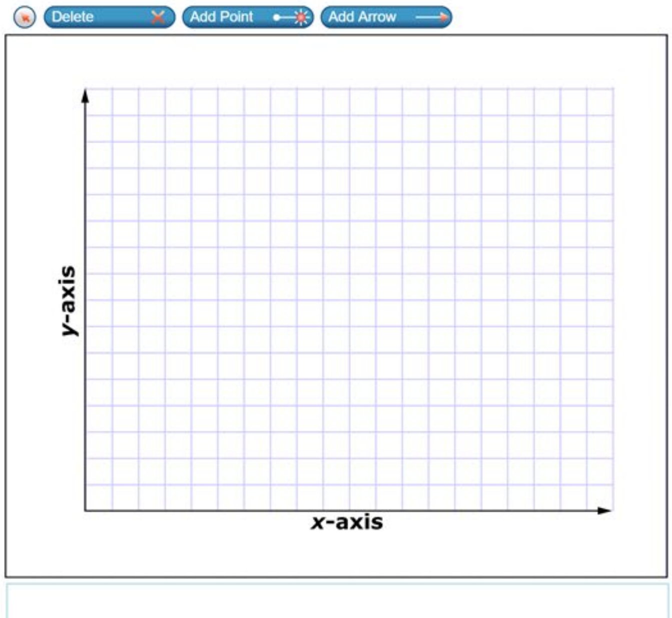
Item 2 of the Blood Sugar Regulation cluster is shown in Figure 35.

Figure 35. Item 2: Blood Sugar Regulation



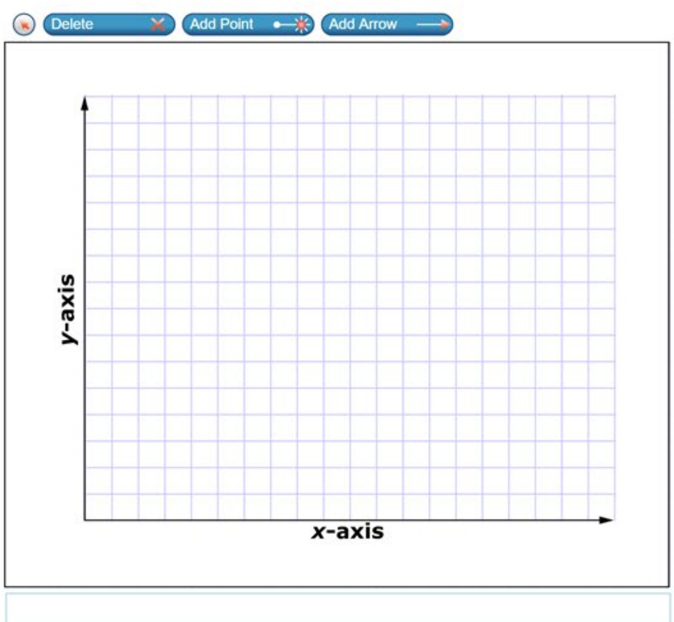
Relationship 2:

x-axis: ▾ *y*-axis: ▾



Relationship 3:

x-axis ▾ *y*-axis ▾



SCORES

Student scores on this item are as follows:

- No students earned 3 score points (full credit).
- Two students earned 2 score points.
- One student earned 1 score point.

COMPREHENSION

Eight students expressed some confusion as to how to construct the graphs of the simulation data. For example, one student was "kind of confused" about where to draw the second and third graphs. Initially she did not see the answer grids for the second and third graphs, but even after she noticed the additional answer grids, some confusion lingered.

At least five students were not sure how to represent the units or values on the graphs, and two students did not draw any graphs for that reason. For example, for the first relationship, one student chose glucose versus time for the first relationship, but he was not sure which value to put on which axis: "I've never looked at the concentration of molecules and tried to graph it, and I feel like there are a lot of things I'm missing to help me figure out what to do. I think I may be overcomplicating it to myself."

REASONING

The following is an example of how one student reasoned through the construction of one of the graphs.

> The student said that he was going to place concentration on the x-axis and time on the y-axis because "in sciences you usually do time on the y-axis and concentration and stuff on the x-axis. I don't know why, it's what I've always known." He selected *Glucose Concentration* for the x-axis and *Time Passed after Eating* for the y-axis. He used the numbers for the glucose concentrations from the simulation in Item 1 to plot points on the graph. He said, "I feel like it spikes up like 5 times so I'll put it a decent amount, 6, 8 and then 10, and it kind of stays pretty high but not as high, so like right there, and then it drops a little bit again, and then it spikes up in a big lunge, and then it drops back down again to here, but it kind of stayed, and then it spiked the highest peak at dinner." He then started to connect the points, and said, "I don't know what the point of the arrows are, I'm just going to connect them all to show their relationship. That's my best guess to show what happened each hour."

Item 3

Item 3 of the Blood Sugar Regulation cluster is shown in Figure 36.

Figure 36. Item 3: Blood Sugar Regulation

> Click on each blank box and select the words or phrases to complete the statements describing the feedback mechanisms that regulate blood sugar levels.
>
> Hunger is part of the ▾ feedback mechanisms, in which the liver and pancreas participate, that ▾ a change in the blood's glucose concentration. The pancreas produces ▾ when blood glucose ▾. The liver responds by ▾ glucose.

Scores

Student scores on this item are as follows:

- Three students earned 2 score points (full credit).
- Seven students earned 1 score point.
- Among these 10 students,
  - four earned a point for correctly filling the blanks in the statement about hunger; and
  - seven earned a point for correctly filling the blanks in the statement about the roles of the pancreas and the liver.

Comprehension

No students expressed confusion about this item.

Reasoning

In responding to Item 3, five students referred to the stimulus, and two students referred to the simulation results in Item 1.

3.4.2 Cluster 2: Saving the Tuna

Performance Summary

The median time to complete the Saving the Tuna cluster was 14 minutes. Table 31 and Table 32 indicate the number of students attaining cluster total scores and items scores within the specified ranges, respectively.

Table 31. Number of Students Attaining Cluster Total Scores in Specified Range: Saving The Tuna

Score 7–6	Score 5–3	Score 2–1	Score 0
1	2	5	4

Note. Maximum score = 7; *n* = 12; three students ran out of time before completing this cluster.

Table 32. Number of Students Attaining Item Scores in Specified Range, by Item: Saving the Tuna

	Maximum Item Score	Score 3	Score 2–1	Score 0
Item 1 (Part A)	3	0	6	6

	Maximum Item Score	Score 1	Score 0
Item 1 (Part B)	1	6	6
Item 1 (Part C)	1	1	11

	Maximum Item Score	Score 2	Score 1	Score 0
Item 2 (Part A and B)	2	3	0	9

Note. n = 12; three students ran out of time before completing this cluster.

Task Demands

The following are task demands of the Saving the Tuna cluster:

- Articulate, describe, illustrate, or select the relationships, interactions, and/or processes to be explained. This may entail sorting relevant from irrelevant information or features.
- Express or complete a causal chain explaining how human activity impacts the environment. This may include indicating directions of causality in an incomplete model such as a flow chart or diagram or completing cause and effect chains.
- Identify evidence supporting the inference of causation that is expressed in a causal chain.

- Use an explanation to predict the environmental outcome given a change in the design of human technology.
- Describe, identify, and/or select information needed to support an explanation.

Stimulus

The stimulus for the Saving the Tuna cluster is shown in Figure 37.

Figure 37. Stimulus: Saving the Tuna

Saving the Tuna

North Atlantic bluefin tuna are one of the most prized fish in danger of overfishing. One 342 kilogram (kg) tuna sold for close to $400,000 dollars at a fish market in Tokyo.

Bluefin tuna are the apex predators in their ecosystem. They hunt, travel, and live within schools, or large groups, of other bluefin tuna individuals. Bluefins start out as extremely tiny larvae, no more than a few millimeters long, and weigh only a few hundredths of a gram. Within three to five years, sexually mature adults can reach lengths of three feet (about one meter) and can weigh over 600 kg. As adults, they can dive as deep as 914 meters and can swim very long distances in the open ocean during migration season. Their migration season spans from approximately May to June, during which they spawn near the Gulf of Mexico.

Because bluefin are prized fish that vary greatly in size and can be found in schools, or groups, within a wide range of water depths, netting fishing methods are commonly used to target and catch these individuals. However, fishing nets often catch bycatch individuals, or non-tuna individuals. The table summarizes several netting fishing methods and the relative amounts of targeted tuna and bycatch individuals caught at one time by each method.

Summary of Netting Fishing Methods

Method	Description	Type of Targetted Catch	Total Number of Individuals Caught at a Time	Percent of Total Catch that is Bycatch (%)	Types of Bycatch Caught
Purse Seining	Large wall of netting that herds fish together and then envelops them when the net is pulled by a drawstring	Schooling or spawning fish	Hundreds to thousands	35 - 70	Sea turtles, dolphins, and other fish
Cast Netting	Small-meshed netting cast from shore or canoes that expands a relatively small area	Groups of small fish	Up to a hundred	10 - 30	Other small fish
Gilnetting	Large curtains of netting suspended by a system of floats and weights that can either be anchored to the seafloor or allowed to float at the surface	All types of fish	Hundreds to thousands	40 - 75	Sea birds, sea turtles, octopi, shark, dolphins, other fish, and crustacea
Midwater Trawling	Gigantic nets that span the size of five football fields pulled by large industrial ships through the open ocean, catching entire schools of fish	All types of open-ocean fish	Thousands to tens of thousands	30 - 75	Sea turtles, shark, dolphins, and other fish
Seine Netting	Small-meshed netting suspended vertically by floats and weights from the surface of intertidal water to enclose and concentrate fish	Crustacea and shell fish	Less than a hundred	10 - 30	Sea birds and other small fish

Your task is to design, evaluate, and refine solutions for reducing the impacts of human fishing on the population of tuna and other native species in the Northern Atlantic Ocean.

Details by Item

Item 1

Item 1 of the Saving the Tuna cluster is shown in Figure 38.

Figure 38. Item 1: Saving the Tuna

The following question has three parts. First, answer part A. Next, answer part B. Then, answer part C.

Part A

Select the boxes to evaluate the tradeoff considerations of each fishing method.

- You may select more than one method per column.

	Likely to Catch the Greatest Number of Tuna Individuals	Likely to Catch the Least Number of Tuna Individuals	Likely to be the Best at Targeting Tuna Individuals	Likely to be the Worst at Targeting Tuna Individuals	Likely to be the Best at Protecting Biodiversity of Ecosystem	Likely to be the Worst at Protecting Biodiversity of Ecosystem
Purse seining	☐	☐	☐	☐	☐	☐
Cast netting	☐	☐	☐	☐	☐	☐
Gilnetting	☐	☐	☐	☐	☐	☐
Midwater trawling	☐	☐	☐	☐	☐	☐
Seine netting	☐	☐	☐	☐	☐	☐

Part B

Based on the evaluation of tradeoff considerations in part A, which fishing method best limits the negative effects of human fishing on non-tuna populations in the Northern Atlantic?

Ⓐ purse seining

Ⓑ cast netting

Ⓒ gilnetting

Ⓓ midwater trawling

Ⓔ seine netting

Part C

Click on each blank box and select a word or phrase to complete a statement describing a change that can be made to decrease the amount of bycatch for the method identified as the worst in targeting tuna individuals in part A.

[▾] the [▾] will improve the targeting of bluefin tuna.

Item 1 (Part A)

Scores

Student scores on this item are as follows:

- No students earned 3 score points (full credit).
- Two students earned 2 score points.
- Four students earned 1 score point.
- Six students earned no score points.

Comprehension

Several students expressed confusion with different aspects of this sub-question including

- completely missing two of the columns in the *Summary of Netting Fishing Methods* table, which was a critical reference for this sub-question; and
- confusion with the response-entry table, including overlooking the instructions stating that it was permissible to select more than one method for each column.

Reasoning

All students methodically navigated through the response-entry table and used the *Summary of Netting Fishing Methods* chart in the stimulus to figure out their responses. For example:

- One student first lined up the *Summary of Netting Fishing Methods* chart next to the response-entry table so that he could read the descriptions easily and fill out the table. For the first column (*Likely to Catch the Greatest Number of Tuna Individuals*), the student said, "The first one I will cancel out will be *cast netting* because it says up to 100, and also *seine netting* because that's less than 100. I would say *gillnetting* and *purse* [are] the two top because it says they catch up to 100s to 1,000s for both of those. Wait; sorry, I was reading that wrong. Okay, *midwater trawling* was 1,000s to 10,000s because that's what I was thinking instead of 100s to 2,000s, so *midwater trawling* will be my answer." The student continued in the same manner for each of the six columns.
- Not all the student's conclusions from the *Summary of Netting Fishing Methods* chart were correct, however, probably because of deficiencies in the student's knowledge about ecology. For example, for column 5 (*Likely to be the Best at Protecting Biodiversity of Ecosystem*), the student said, "I would say both *gillnetting* and *midwater trawling* because they both take all types of fish, they are not going after specific fish, which means that they're not taking one species of fish out of the water; they're taking multiple, so there's less chance of one fish being taken out of the ecosystem."

Item 1 (Part B)

Scores

Six students earned credit on this sub-item.

Comprehension

One student was confused, saying that she did not understand the question and she did not know about each type of net.

Reasoning

In responding to this sub-item, four students referred to their responses in Part A, and four students referred to the *Summary of Netting Fishing Methods* chart.

Item 1 (Part C)

Scores

One student earned credit on this sub-item.

Comprehension

Several students clearly did not understand the sub-item and guessed on questionable grounds.

> For example, one student read out loud all of the options under the second drop-down menu and said that he did not really understand the question: "I'm confused because in re-reading the question, it makes it seems like it was asking which net would decrease the chance of getting a tuna, but re-reading the answer choices, it's not asking that as much as I thought it would be. So, I'm going to go with *decreasing* instead of *increasing* because it says decrease in the sentence, and then something about negatives."

Another student indicated that she initially thought the sub-item was looking for a change in any of the methods that would decrease the amount of tuna by catch. Later she realized that the sub-item was referencing something specific in Part A. She went through all the drop-down options and hesitated a lot over her answer, changing it several times.

Reasoning

In responding to this sub-item, five students referred to their responses in Part A, and six students referred to the *Summary of Netting Fishing Methods* chart.

Item 2

Item 2 of the Saving the Tuna cluster is shown in Figure 39.

Figure 39. Item 2: Saving the Tuna

The following question has two parts. First answer part A. Then, answer part B.

Three solutions proposed by scientific and environmental organizations to protect and restore the Northern Atlantic bluefin tuna population are shown in the table.

Solutions to Protect and Restore the Bluefin Tuna Populations

Solution	Description
1	Completely restricting the catching of juvenile bluefin
2	Limiting the total number of adult bluefin that can be caught
3	Removing juvenile bluefin from the Northern Atlantic to raise in captivity

Part A

Which Bluefin characteristic serves as the criteria on which all three solutions are based?

Ⓐ body mass

Ⓑ body length

Ⓒ ability to reproduce

Ⓓ ability to dive for prey

Part B

Select the **two** netting characteristics that are most important to consider when designing fishing nets for use in implementing the three solutions.

☐ mesh size of the net

☐ overall size of the net

☐ ability of the net to move

☐ depth of the net’s location within the water column

Scores

Student scores on this item are as follows:

- Three students earned 2 score points (full credit).
- No students earned 1 score point.
- Nine students earned no score points.

- Part A contributed one-third of the weight to the total item score, and 11 students selected the correct response for Part A.
- Part B contributed two-thirds of the weight to the total item score. Students only received credit for Part B if they correctly identified two netting characteristics that are important to consider when designing fishing nets for use in implementing the three solutions. While only three students correctly selected both characteristics, seven other students correctly selected one of the characteristics (four selected the *depth of the net's location in the water* column, and three selected the *mesh size of the net* column).

COMPREHENSION

One student did not understand the term "mesh size." She understood mesh as a verb, e.g., "meshing things together."

REASONING

When responding to Part B, only one student referred to the *Solutions to Protect and Restore the Bluefin Tuna Populations* table included with the item; four students referred to the *Summary of Netting Fishing Methods* chart in the cluster stimulus, and two students referred to the text in the cluster stimulus.

The following is an example of how one student used the reference materials to draw two conclusions about how to design the net to protect and restore the tuna population. Rather than considering any of the solution strategies proposed in the cluster stimulus, the student seemed to focus on supporting a method that would selectively catch adult tuna rather than juveniles, but one of the net characteristics he identified (*depth of the net's location within the water colum*n) counted as correct.

> The student looked at the fishing method characteristics and said, "They're going to want to increase the depth of the net's location within the water column because the adults can dive as deep as 914 meters and can swim very long distances, so they're going to want to increase the depth and the overall size of the net to catch them." When asked where the student got the information to answer the question, the student said, "I looked at the top of the article where it says that they dive as deep as 914 meters and can swim very long distances in the open ocean. So, I said increase the overall size to make the catch wider so they can't swim outside of the range of the net and also increase the depth since they can go pretty low."

3.4.3 Cluster 3: Tomcods

Performance Summary

The median time to complete the Tomcods cluster was 17 minutes. Table 33 and Table 34 indicate the number of students attaining cluster total scores and item scores within the specified ranges, respectively.

Table 33. Number of Students Attaining Cluster Total Scores in Specified Range: Tomcods

Score 8–6	Score 5–4	Score 3–1	Score 0
0	1	9	4

Note. Maximum score = 8; *n* = 14; one student ran out of time before completing this cluster.

Table 34. Number of Students Achieving Item Scores in Specified Range, by Item: Tomcods

	Maximum Item Score	Score 5–4	Score 3–1	Score 0
Item 1 (Parts A–C)	5	0	2	12

	Maximum Item Score	Score 1	Score 0
Item 2 (Part A)	1	6	8
Item 2 (Part B)	1	0	14
Item 3	1	10	4

Note. n = 14; one student ran out of time before completing this cluster.

Task Demands

The following are task demands of the Tomcods cluster:

- Based on the provided data, identify, describe, or construct a claim regarding the effect of changes to the environment on (1) the increases in the number of individuals of some species, (2) the emergence of new species over time, and (3) the extinction of other species.
- Sort inferences about the effect of changes to the environment on (1) the increases in the number of individuals of some species, (2) the emergence of new species over time, and (3) the extinction of other species into those that are supported by the data, contradicted by the data, outliers in the data, or neither, or some similar classification.
- Identify patterns of information/evidence in the data that support correlative/causative inferences about the effect of changes to the environment on (1) the increases in the number of individuals of some species, (2) the emergence of new species over time, and (3) the extinction of other species.

- Construct an argument using scientific reasoning drawing on credible evidence to explain the effect of changes to the environment on (1) the increases in the number of individuals of some species, (2) the emergence of new species over time, and (3) the extinction of other species.
- Identify additional evidence that would help clarify, support, or contradict a claim or causal argument regarding the effect of changes to the environment on (1) the increases in the number of individuals of some species, (2) the emergence of new species over time, and (3) the extinction of other species.
- Identify, summarize, or organize given data or other information to support or refute a claim regarding the effect of changes to the environment on (1) the increases in the number of individuals of some species, (2) the emergence of new species over time, and (3) the extinction of other species.

Stimulus

The stimulus for the Tomcods cluster is shown in Figure 40.

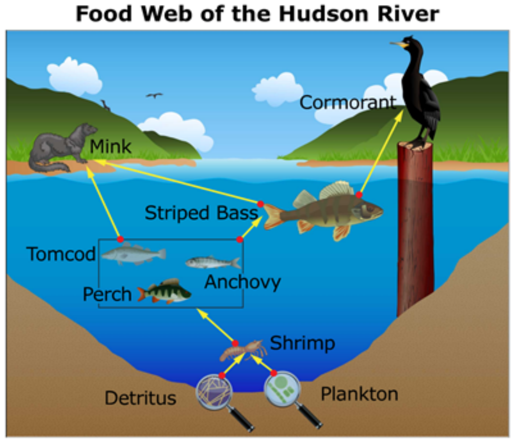
Figure 40. Stimulus: Tomcods

Atlantic Tomcod Thrive in Contaminated Hudson River

Polychlorinated biphenyls (PCBs) are chemicals that were produced from 1929 to 1979 for industrial and commercial uses. One electric company released 1.3 million pounds of PCBs into the Hudson River from 1947 to 1976. In 1979, PCBs were banned. However, the Hudson River still has high levels of PCBs today because they settle into sediments on the bottom and do not break down. When most fish embryos are exposed to PCBs, the immune system of the embryo is disrupted, causing the fish to develop smaller hearts that do not function properly, resulting in death. Many fish populations declined or disappeared from the Hudson River because of PCB exposure. However, one fish population, the Atlantic Tomcod, does not have this reaction to PCBs and thrives.

The picture shows a food web for the Hudson River. The liver of several aquatic species were tested for the presence of PCBs. The levels of PCBs in the livers of the tomcod were among the highest reported. Both striped bass and mink populations have also been found to have high levels of PCBs.

Food Web of the Hudson River



Tomcod were captured from the Hudson River and from rivers not contaminated by PCBs. The tomcod were tested for the AHR2 protein, which is responsible for regulating the toxic effects of PCB. The percentage of tomcod that contained the AHR2 protein mutation is shown in the table.

Percentage of Tomcod with AHR2 Protein Mutation

River	Percentage of Tomcod with Mutation
Hudson River, New York	99
Hackensack River, New Jersey	92
Niantic River, Connecticut	6
Shinnecock Bay, New York	5

Following are two hypotheses about the success of the tomcod in the contaminated Hudson River.

Hypothesis 1: The tomcod population did not decrease in response to PCB exposure because tomcod do not take in as many PCBs as other fish species through their food consumption or absorption from the water.

Hypothesis 2: The tomcod population did not decrease in response to PCB exposure because they have evolved resistance to the effects of PCBs through natural selection.

As you work through the questions, evaluate the evidence to determine which hypothesis of how the tomcods are able to overcome exposure to deadly PCBs is **best** supported.

Reference: Isaac Wirgin, et al. "....Atlantic Tomcod from the Hudson River." *Science* 331 (2011):1322–1325.

Details by Item

Item 1

Item 1 of the Tomcods cluster is shown in Figure 41.

Figure 41. Item 1: Tomcods

The following question has three parts. First, answer Part A. Next, answer part B. Then, answer part C.

Part A

Select the boxes to indicate whether each statement supports or refutes Hypothesis 1 or Hypothesis 2. You can select more than one box for each statement.

	Supports Hypothesis 1	Refutes Hypothesis 1	Supports Hypothesis 2	Refutes Hypothesis 2
There is a higher percentage of AHR2 protein mutations in the Hudson River than in rivers not contaminated by PCBs.	☐	☐	☐	☐
PCBs accumulate in striped bass and mink as a result of food consumption.	☐	☐	☐	☐
There is a high level of PCBs in the liver of tomcod in the Hudson River.	☐	☐	☐	☐
The tomcod population thrives in the PCB-contaminated Hudson River.	☐	☐	☐	☐
Tomcod feed on small PCB-contaminated bottom feeders but do not show any effects of PCB-exposure.	☐	☐	☐	☐

Part B

Click on each box to select the word or phrase that **best** completes the statement.

[▾] is most probable because [▾] the evidence supports this hypothesis and [▾] the evidence refutes this hypothesis.

Part C

Select additional evidence to support the hypothesis selected in part B.

☐ The Hudson River shrimp and plankton do not take in as much PCB as the fish species.

☐ DNA evidence shows changes to the gene for AHR2 in the tomcod of the Hudson River.

☐ Changes to the AHR2 protein are acquired in response to environmental cues and are not genetic.

☐ The Hackensack River shares an estuary with the Hudson River, allowing fish to pass genes back and forth.

SCORES

Student scores on this item are as follows:

- No students earned 5 score points (full credit) on this item.
- The highest score earned was 2 points, and this was achieved by two students, who each earned 1 point for Part A and 1 point for Part B. No one achieved any points for Part C.
- The remaining 12 students earned no credit.

COMPREHENSION

It is hard to extract any detailed information on students' comprehension or reasoning because students floundered so badly on this question.

REASONING

In Part A, most students did conscientiously work their way through the list of evidence and try to determine which supported or refuted each hypothesis, but their reasoning was substantially flawed, perhaps because they did not understand the applicable content knowledge.

> For example, one student read out loud Hypothesis 1 and 2 in the introduction. She said, "So there's a higher percentage in the Hudson River than in rivers not contaminated," and selected Supports Hypothesis 1 for line 1 "because it's talking about how this one is saying that it's from the water and not from the fish." She read out loud part of line 2, looked quickly at the table in the introduction, and said that it's "actually going against it [refutes Hypothesis] because this one is talking about how it's because of the water not because of the fish, because of the food they are consuming, and they are not talking about the actual fish," then clicked Refutes Hypothesis 1. She read out loud line 3. She said she was going to select Refutes Hypothesis 1 because "it's the same as the first one, because it's saying how the species through the food, not the fish itself." She read out loud line 4 and immediately said that it supports Hypothesis 2 because "it's talking about how it is contained in the actual river, not the fish's fault, but the river's fault." She read out loud line 5 and said immediately that line 5 also supports Hypothesis 2 because, "of the natural selection."

Students who did not have good comprehension of Part A had even less chance of reasoning their way through Parts B or C, both of which built on conclusions from Part A.

Item 2

Item 2 of the Tomcods cluster is shown in Figure 42.

Figure 42. Item 2: Tomcods

The following question has two parts. First, answer part A. Then, answer part B.

Part A

Why were the tomcod able to survive in the presence of PCBs when other species were not?

Ⓐ The Hudson River tomcod did not absorb PCBs from the water.

Ⓑ All populations of tomcod species are resistant to the effects of PCB.

Ⓒ The Hudson River tomcod did not feed on species that were contaminated with PCBs.

Ⓓ The AHR2 mutation already existed in the Hudson River tomcod population at a low frequency.

Part B

Select the evidence that supports your answer.

☐ All tomcod tested in all rivers were resistant to PCB exposure.

☐ None of the Hudson River tomcod were found to contain PCBs.

☐ The AHR2 protein mutation is found at low frequency in tomcod from rivers not contaminated with PCBs.

☐ Less than 50 years after first exposure to PCBs, almost all of the Hudson River tomcod could survive in the presence of PCBs.

SCORES

Student scores on this item are as follows:

- Six students earned credit on Part A by choosing the correct explanation for why Tomcods can survive in the presence of PCBs.
- Three of those students also selected one of the pieces of evidence that supported their explanation, but they received no credit for Part B because they did not select both the applicable pieces of evidence.
- Three other students also selected one piece of "correct" evidence, but they had not chosen the right explanation in Part A, so it was unclear exactly what they were supporting.

COMPREHENSION

Although it was hardly the only reason why students had difficulty with this item, students were clearly challenged by having to pick more than one right answer in Part B, perhaps because they are not familiar with multi-select items and just stopped looking after they had made one selection. It might have helped to cue the students if the stem had specified that they had to select ALL the evidence that supported their explanation.

REASONING

The following is an example of the reasoning of one of the students who correctly identified option D as the reason why Tomcod survived in Part A,

> The student read option A out loud and said, "That's a lie! Because it says up there tomcod have a bunch of it, so that's definitely a lie." The student read option B out loud, saying, "I'm going to say No, because, in the [student looked back to the table on the left] Niantic River and the Shinnecock Bay, they did not have that mutation. So, I'm going to say B is wrong." The student read option C out loud, saying, "OK wrong, because they eat the plankton and the shrimp, and they said earlier that they eat bottom feeders that have it." Student read option D out loud and said, "Yes, because then they would have made it and had a bunch with that mutation."

Item 3

Item 3 of the Tomcods cluster is shown in Figure 43.

Figure 43. Item 3: Tomcods

Why were other fish species in the Hudson River wiped out by PCB exposure, while the tomcod thrived?

- (A) Other species do not contain a protein that regulates the toxic effects of PCBs, so they could not adapt quickly.
- (B) Other species consumed more contaminated food than the tomcod, so they had more severe effects from PCB exposure.
- (C) Other species absorbed the PCBs from the water more quickly than the tomcod, so they had higher concentrations in their bodies.
- (D) Other species could not adapt quickly because they did not already contain a beneficial mutation in the gene pool to protect them from the effects of PCBs.

SCORES

Students did the best on this item; 10 students earned credit.

COMPREHENSION

No features of this item appeared to confuse students.

REASONING

Students who chose the right answer demonstrated plausible reasoning that supported the inference that the students had mastered the concept being tested.

> For example, one student read out loud response option A and said, "That's a good one, that might be the one." He read out loud response option B and said, "That one does not make any sense because all fish, I'm assuming. [are] about the same size will eat about the same, and I know that goldfish don't fill their stomach. I believe they go for all fish, they are all eating like crazy, so I would not click that one." He read out loud response option C twice and said, "Again, that's the same explanation for C as B, I would not click it." He

read out loud response option D and said, "That's the one I'm going to click, because that one is exactly referring to natural selection and . . . it's like a gene, something in their mutation that they could protect themselves from the effects of it, but it's in the gene pool and it's referring to natural selection and the crossing of two species to get your genes and I would go with D, and A would be a close choice."

3.4.4 Cluster 4: Tuberculosis

Performance Summary

The median time to complete the Tuberculosis cluster was 10 minutes. Table 35 and Table 36 indicate the number of students attaining cluster total scores and items scores within the specified ranges, respectively.

Table 35. Number of Students Attaining Cluster Total Scores in Specified Range: Tuberculosis

Score 5–4	Score 3–1	Score 0
1	9	4

Note. Maximum score = 5; *n* = 14; one student ran out of time before completing this cluster.

Table 36. Number of Students Attaining Item Scores in Specified Range, by Item: Tuberculosis

	Maximum Item Score	Score 3	Score 2–1	Score 0
Item 1	3	1	5	8

	Maximum Item Score	Score 1	Score 0
Item 2 (Part A)	1	6	8
Item 2 (Part B)	1	1	13

Note. n = 14; one student ran out of time before completing this cluster.

Task Demands

The following are task demands of the Tuberculosis cluster:

- Based on the provided data, make or construct a claim regarding inheritable genetic variations that may result from: (1) new genetic combinations through meiosis, (2) viable errors occurring during replication, and/or (3) mutations caused by environmental factors. This does not include selecting a claim from a list.
- Sort inferences about inheritable genetic variation into those that are supported by the data, contradicted by the data, outliers in the data, or neither, or some similar classification.
- Identify patterns of information/evidence in the data that support correlative/causative inferences about inheritable genetic variation.
- Construct an argument using scientific reasoning drawing on credible evidence to explain inheritable genetic variations may result from: (1) new genetic combinations through meiosis, (2) viable errors occurring during replication, and/or (3) mutations caused by environmental factors (handscored constructed response).

- Identify additional evidence that would help clarify, support, or contradict a claim or causal argument.
- Identify, describe, and/or construct alternate explanations or claims and cite the data needed to distinguish among them.
- Predict outcomes of genetic variations, given the cause and effect relationships of inheritance.

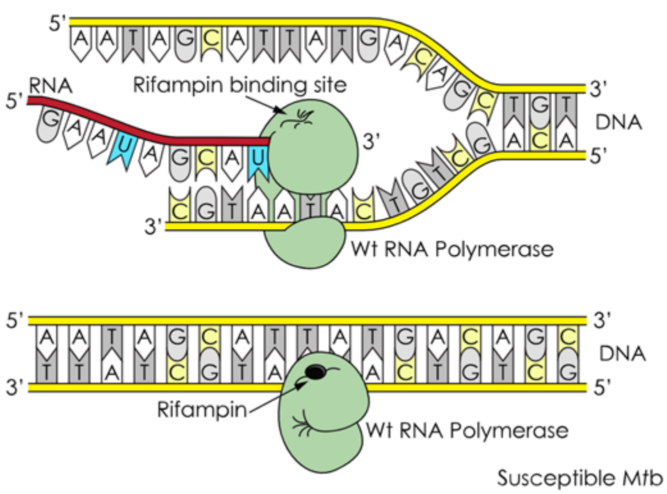
Stimulus

The stimulus for the Tuberculosis cluster is shown in Figure 44.

Figure 44. Stimulus: Tuberculosis

Antibiotic Resistant Tuberculosis

Antibiotic-resistant bacteria present a growing health care problem. The bacteria *Mycobacterium tuberculosis (Mtb)* causes the disease tuberculosis. One antibiotic used to treat tuberculosis is rifampin. Rifampin works by binding to amino acids 36-67 of the RNA polymerase protein of *Mycobacterium tuberculosis.* This binding makes the RNA polymerase protein inactive and the cell dies. This is illustrated below:



However, when treated with the antibiotic rifampin, some *Mycobacterium tuberculosis* bacteria are killed, but others survive. The bacteria that are killed are called "susceptible" to the antibiotic.

Scientists grow 3 mutant strains of *Mycobacterium tuberculosis* bacteria in a lab and sequence their DNA to compare to the wild-type strain that is not resistant to rifampin. Review the information provided.

Comparison of Mutant *Mycobacterium Tuberculosis* Bacteria to Wild-Type

Strain	DNA Sequence Change	Amino Acid Position	Amino Acid Change
Mutant 1	G to A substitution mutation	30	Alanine to Threonine
Mutant 2	C to A substitution mutation	51	No change
Mutant 3	G to T substitution mutation	46	Aspartic Acid to Tyrosine

As you work through the questions, evaluate the evidence to identify the source of genetic variation for antibiotic resistance in *Mycobacterium tuberculosis*.

Details by Item

Item 1

Item 1 in the Tuberculosis cluster is shown in Figure 45.

Figure 45. Item 1: Tuberculosis

If the rifampin cannot bind to the RNA polymerase protein in *Mycobacterium tuberculosis*, this leads to antibiotic resistance. Mutations in the rifampin binding site can block binding of the antibiotic. Based on the information provided, determine which mutants are likely to be resistant to rifampin by this mechanism.

Click on each blank box to select the correct words or phrases.

Resistance of Mutant *Mycobacterium Tuberculosis* Strains

Strain	Resistance	Explanation		
Mutant 1	▾	▾	▾	of rifampin
Mutant 2	▾	▾	▾	of rifampin
Mutant 3	▾	▾	▾	of rifampin

Scores

One student earned 3 score points (full credit), and she was the only one to earn a point for correctly determining and explaining the resistance status of Mutant 3.

Five other students each earned 1 score point. Three of these students earned their point for correctly determining and explaining the resistance status of Mutant 2, and two earned their point for Mutant 1.

Comprehension

Four students reported that they found this item confusing and did not understand how to derive the necessary information from the stimulus.

> For example, one student said that Item 1 was confusing and that it was not really addressed [in the stimulus]. He said he was doing a lot of "assuming" because "it's talking about 'resistant,' and he only saw the word once." He also said that "it seemed weird that all three of them would be not resistant," although it is not clear on what basis he concluded that all three mutant strains were not resistant.

Four students reported using things they learned in science classes at school to help them respond to this item. For example,

- one student said that she knew about the amino acid from Biology in freshman year, and
- another student said that he learned about the topic in a biotech class two weeks prior to the interview.

REASONING

All but two of the students referred to the comparison table in the stimulus when responding to this item; four students referred to the diagram.

Although only one student had the correct responses for all three of the mutant strains, several used the stimulus materials in the intended manner to reason through the problem.

> For example, one student looked at the comparison table in the stimulus and said, "It says that the Rifampin works by binding to amino acids 36-67 of the RNA. And then it says down here that, because of the G to A substitution mutation, the amino acid positions at number 30, and then . . . it is resistant because it changed it from 36 to 30, so then the Rifampin can't bind to it…So I would say it's resistant, but there's no change of rifampin—oh yeah, change to the—outside of the binding site." "Mutant 2 changed it C to A. Mutant 2 changes the amino acid to 51, so there's no change, so I'm going to mark *Not Resistant* because it's still within 36-67, so I'm going to say no change inside the binding site." "And Mutant 3 is a G to T substitution to 46. And 46 is still within 36-67, so I'm going to say *Not Resistant*, because there is a change from aspartic acid to tyrosine, Inside the binding site."

Item 2

Item 2 of the Tuberculosis cluster is shown in Figure 46.

Figure 46. Item 2: Tuberculosis

The following question has two parts. First, answer part A. Then, answer part B.

Part A

What is the **likely** source of the genetic variation in antibiotic resistance of *Mycobacterium tuberculosis*?

Ⓐ new genetic combinations through meiosis

Ⓑ new genetic combinations through mitosis

Ⓒ viable errors occurring during DNA replication

Ⓓ sexual reproduction resulting in new combinations of traits

Part B

From the list of additional experiments, select the evidence that would support your answer in part A.

☐ Scientists grow a sample of wild-type *Mycobacterium tuberculosis* in the lab. Over time, some of the bacteria show resistance to rifampin.

☐ Scientists plate a colony of wild-type *Mycobacterium tuberculosis* and a colony of *Escherichia coli* in one petri dish. Some of the new colonies show resistance to rifampin.

☐ Scientists plate a colony of wild-type *Mycobacterium tuberculosis* and a colony of mutant *Mycobacterium tuberculosis* in one petri dish. Some of the new colonies show resistance to rifampin.

☐ Scientists create additional *Mycobacterium tuberculosis* mutants by creating substitution mutations in the DNA that codes for amino acids 36-67. Many of the mutants are resistant to rifampin.

Item 2 (Part A)

Scores

Half of the students (seven students) earned credit on this sub-item.

Comprehension

No features of this item appeared to confuse students.

Reasoning

Three students looked back to one or more parts of the stimulus while working on this sub-item.

Four students said they used, or tried to use, material learned in school to help them respond to this sub-item. For example,

- one student said, “I am trying to go back to my knowledge of mitosis and meiosis and DNA replications,” and

- another student said, "Usually errors that occur during DNA replication can be bad, and I remember back from when I was a freshman that it's not hereditary."

Some students used test-wise strategies to make plausible guesses, so a correct answer did not necessarily represent full mastery.

> For example, one student (who correctly selected C, *viable errors occurring during DNA replication*) said in his think aloud, "All this right now has to do with DNA . . .I don't see anything about meiosis and mitosis on the chart." When asked how he came up with his answer, he said, "I didn't think it was A or B cause it's talking about meiosis and mitosis, which was not discussed in the article, and then same with D. I did the viable errors because it's talking about DNA strands, so that's why I chose C."

Item 2 (Part B)

SCORES

Only one student earned credit for this sub-item. In part, the difficulty resulted from an incorrect interpretation of the sub-item, as explained further in the Comprehension section below.

Of the two correct options, five students selected *Scientists grow a sample of wild-type Mycobacterium tuberculosis in the lab* . . . and seven students selected *Scientists create additional Mycobacterium tuberculosis mutants by creating substitution mutations in the DNA* . . .

COMPREHENSION

To earn credit for this item, students had to select both the experiments that could provide evidence to support the conclusion they selected in Part A. However, this is not clearly stated in the instructions, so most students stopped after they thought they had found one relevant experiment. Only three students marked two options, and two students said that they thought that they were only allowed to choose one option.

One student expressed confusion with the second response option. He did not know what *Escherichia coli* was and the relationship might be between it and *Mycobacterium tuberculosis*.

REASONING

At least four students referred to the text, diagram, and/or comparison table when responding to this sub-item.

3.5 Students' Overall Perceptions of the Test

3.5.1 Topics Studied

Elementary School (n=18)

- Eleven students reported that they had studied topics related to the Desert Plants cluster, such as the life cycle of a plant and how plants survive in a desert habitat.
- Ten students had studied topics related to the Grand Canyon cluster, although not all of them learned about fossils or contemporary animals that can be found in the canyon. One student learned about fossils and rock formations as part of the history of Utah.
- Nine students had studied topics related to the Terrarium Matter Cycle cluster, such as "plants have carbon dioxide, but a whole plant needs water, soil, and sun," and some had conducted an experiment in which one group of students tried to grow plants in a dark environment and another group tried to grow plants in the sunlight.
- Although no students were familiar with topics related to the German Pyramid Candle cluster, five students had studied heat transfer.

Generally, each of the Utah students had studied more of these topics than the California students, and their lessons were more closely aligned with the topics of the science clusters. One of the Utah students said he had studied all four of the topics:

> "At the beginning of the year we studied the heat one and how we can help make a motor turn something on, like a light bulb. I thought of that. Maybe it was just backwards, the light was helping the fan to spin. The light was turning or making it spin by the energy it was producing. I remember last year in 4th grade we studied the Grand Canyon and the animals, and we did a little bit this year, and the animals that were living in the walls like trilobite and some others like starfish. We saw this video of this hole that was in Arizona, and there were tons of fossils in it. I think we studied a little bit on the terrarium one . . . We studied a little bit about [the desert plants]. About how each plant could survive."

Middle School (n = 12)

- Nine of the 11 students who responded to the Galilean Moons cluster question reported that they had studied related topics, such as moons, the solar system, space, and the planets, although their studies were not as in-depth as the animation and the data table.
- Only three students had studied the water cycle or how it applied to fog.
- Four students had studied some aspects of weather, including warm and cold fronts, but not as in-depth as the Texas Weather cluster.
- Eight students had studied animals and the types of relationships between animals, although not necessarily about hippos.

High School (n = 15)

- Thirteen students reported that they had studied topics related to the Tuberculosis cluster, such as DNA, mutations, mitosis, meiosis, and amino acids.
- Seven students had studied topics related to the Blood Sugar Regulation cluster, although not as in-depth as these questions. In referring to the Blood Sugar Regulation cluster, one student said that they had reviewed molecule concentrations but never discussed meals or "not that in-depth, more gone over these and what they do for the body." Another student said she had studied feedback loops and homeostasis.
- Five students had studied topics related to the Tomcods cluster, such as the food web, ecology, and PCBs.
- Only two students said that they had studied topics related to the Saving the Tuna cluster, but they did not provide any information about which specific topics.

3.5.2 Use of Similar Online Tests and Tools

Elementary School (n=18)

All but one student had previously taken online tests; the subjects of the tests varied and included science, mathematics, reading, and/or "grammar." The online tests they had used included Galileo, SALT, ATI, and, for the Utah students, SAGE.

All but one of the students said that they had used similar online tools, including being able to expand the screen from left to right and vice versa; videos; dictionaries; navigation buttons such as arrows, a scroll bar, Back, Next, and Zoom in/Zoom out buttons; and drop-down menus. One student said that her previous experience with online tests involved individual questions rather than clusters, and another student said that there were "more pictures to move around" on the other online test.

Middle School (n = 12)

All 11 students who responded to this question had previously taken online tests; the subjects varied and included science, mathematics, and/or English language arts.

All but two of the students said that they had used similar online tools (including the Connect Line tool and Graphing tool for plotting points), animations, videos, and navigation buttons such as the Next, Back, Pause, and Zoom in/Zoom out buttons. One student said that he previously had to draw lines, but only straight lines, nothing like the graphs she had to draw in the Morning Fog cluster. Another student mentioned that layout of the items was familiar, including having the stimulus on the left side of the screen and the questions on the right side.

High School (n = 15)

All but two students had previously taken online tests; the test subjects varied and included science, mathematics, and English.

All but one of the students said that they had used similar online tools including at least one of the following: graphs, diagrams, the Connect Line tool, checkboxes, and a layout that presented a stimulus on one side of the screen and the associated questions on the other side. One student said that a standardized test he took the previous day was exactly the same, "the interface is the same," although he was not able to expand the screen on the standardized test. One student mentioned two other functionalities that he had used on other tests: the Highlighting tool and the ability to add a note to a paragraph and view it later.

3.6 Overall Thoughts about Test Difficulty

Elementary School (n=18)

Nine students felt that the test had both easy and hard parts and described the overall difficulty as "in between." Examples include the following:

- One student said, "I think the test was in between those because some of it I got confused on and some other pieces like this [referring to Item 1 of the Redwall Limestone cluster] was easy since it gave us these maps about where it lived and the rest was kind of simple. For this one [referring to Item 2 of the Redwall Limestone cluster], it was simple."
- One student said, "Some of them were hard, some of them were confusing, some of them were easy – that's how I feel about this test. The hardest part was [the Terrarium Matter Cycle cluster], question two, Part A [of the Terrarium Matter Cycle cluster] because "I didn't understand what they meant about X, Y, and Z – I had to think about what they mean."
- Another student thought the test was "right in the middle, good. It wasn't too easy or too difficult." The student did not find any of it particularly confusing.
- Five students described only one of the items as being difficult, and four of the five students said the hard item was Item 2 Part A in the Terrarium Matter Cycle cluster. Examples include the following:
    - One student said, "There was one I skipped. I didn't really like that. Because there was too much going on," referring to Item 2 in the Terrarium Matter Cycle cluster.
    - One student felt that the hardest question was on "the terrarium with the diagram and the X, Y, and Z stuff. The others you just had to think about, and you could solve them."
    - Another student said, "Overall, I think it's really good. I found the terrarium a little confusing. It is a good test to have about things you need to know." When asked if the questions were hard or easy, the student said they were easy except for the terrarium question. He said he got confused on the circle of energy.

By contrast, four students expressed that the test was easy. Examples include the following:

- One student did not feel like any of it was confusing, and he was not nervous. He thought the questions were very specific. It was easy for him to navigate through the tools and figure out how to answer the questions.
- One student said, "It took some time for me to think of the answers, but I thought it was pretty easy."

Middle School (n = 12)

All 12 students responded to the end-of-test question on what they thought of the test. Seven of the students felt that the test was not too hard. For example:

- One student thought that the questions were reasonably easy but were hard for someone who hadn't learned a lot of this material. She said that, in general, she is well educated in science, but a lot of these topics are "very random." The student felt like she could have told the interviewer about the water cycle, but not how it works in this specific scenario.
- One student said that the test "was good, yeah. It wasn't hard." The student said that Item 3 of the Galilean Moon cluster was hard.
- Another student thought the questions got harder as she went along, and the hardest problem was the Texas Weather cluster. She had to reread some of the questions, but overall, she thought they were clear.

By contrast, five students expressed that the test was difficult or challenging. For example:

- One student thought that the test was good, but kind of difficult. She mentioned that students like her brother, who is dyslexic, would find it helpful to have the questions read out loud to them. She also said some of the questions were harder because she hadn't gone over the content yet and didn't know what some of the moons were.
- Another student thought the test was "pretty difficult." It was confusing for the student because she had to go back and reread items to understand the process and how to figure it out.
- A student said it was definitely "more challenging" than tests he had taken.
- A student said, "I thought it was kind of confusing. We've studied the moon one a bit, the hippos for sure, and then the water cycle and the temperature we haven't, so for doing all of those for my first time, I couldn't quite make it out. I was totally lost on the Morning Fog in the Valley."

High School (n = 15)

All 15 students responded to the end of the test question on what they thought of the test, although three students did not comment on whether the test was easy or difficult. (One of these latter students described it as "pretty interesting" and "different." Another said he liked the multiple-choice items, the diagrams, tables, and having multiple parts to a question.)

Ten students felt that the test was in the "middle range" of difficulty, with some questions being clearer than others. Four students felt that the Tomcods cluster was confusing, and three students felt that the Blood Sugar Regulation cluster was confusing.

Two students described the test as being difficult. One of these students said the test did not relate to his past studies, but he thought it would be a good test for students who were studying these topics. He also said the types of questions were different than he was used to: – "it's not like normal standardized testing kinds of questions." The student noted that he had not studied these topics even though he was an Advanced Placement (AP) Biology student. Consequently, he was unsure who the target audience of the test might be. The other student mentioned that she found the questions "kinda hard" because there were so many parts to each question. The reading parts were clear, but the structure of the questions could be confusing, according to the student.

Appendix 1: Characteristics of Sample, by Cluster Grade Level and Student

Table 1-A. Elementary School Sample

Student	Location	Grade	Gender	Lunch Program	Ethnicity	Language at Home	IEP (Disability)	Science Grades
1	California	5	Male	No	Asian	English	No (N/A)	Mostly A's
2	California	5	Male	No	Caucasian	English	No (N/A)	Mostly A's
3	California	5	Male	No	Asian	English	No (N/A)	Mostly A's
4	California	5	Male	No	Caucasian	English	No (N/A)	Mostly A's
5	California	5	Male	No	African American	English	No (N/A)	Mostly B's
6	California	5	Male	No	Caucasian	English	No (N/A)	Mostly A's
7	California	5	Female	Yes	Other	English	No (N/A)	Mostly B's
8	California	5	Male	Yes	Caucasian	English	No (N/A)	Mostly A's
9	California	5	Male	Yes	Hispanic	English	No (N/A)	Mostly A's
10	California	5	Male	No	Caucasian	English	No (N/A)	Mostly B's
11	California	5	Female	No	Caucasian	English	No (N/A)	Mostly B's
12	California	5	Female	No	Caucasian	English	No (N/A)	Mostly B's
13	Utah	6	Male	–	Caucasian	–	–	–
14	Utah	6	Male	–	Caucasian	–	–	–
15	Utah	5	Male	–	Caucasian	–	–	–
16	Utah	6	Female	–	Caucasian	–	–	–
17	Utah	5	Male	–	Caucasian	–	–	–
18	Utah	5	Female	–	Caucasian	–	–	–

Note. –: Missing data

Table 1-B. Middle School Sample

Student	Location	Grade	Gender	Lunch Program	Ethnicity	Language at Home	IEP (Disability)	Honors/ Advanced Classes	Science Grades
1	California	9	Female	No	Other	English	No (N/A)	Math	Mostly A’s
2	California	9	Male	No	African American	English	No (N/A)	None	Mostly B’s
3	California	9	Female	No	Caucasian	English	No (N/A)	None	Mostly A’s
4	California	8	Female	No	Caucasian	N/A	No (N/A)	None	Mostly A’s
5	California	9	Female	No	Asian	English	No (N/A)	Math, Science, Reading	Mostly A’s
6	California	8	Female	No	Caucasian	English	No (N/A)	Math	Mostly A’s
7	California	9	Male	Yes	Caucasian	English	Yes (Specific Learning Disability)	None	Mostly A’s
8	California	8	Male	Yes	Hispanic	English	No (N/A)	None	Mostly A’s
9	California	8	Male	Yes	Caucasian	English	No (N/A)	None	Mostly A’s
10	California	8	Male	No	African American	English	No (N/A)	None	Mostly A’s
11	California	8	Male	No	Asian	English	No (N/A)	Math, Science, Reading	Mostly A’s
12	California	8	Female	No	Asian	English	No (N/A)	None	Mostly A’s

Table 1-C. High School Sample

Student	Location	Grade	Gender	Lunch Program	Ethnicity	Language at Home	IEP (Disability)	Honors/ Advanced Classes	Science Grades/ Achievement*
1	California	11	Female	No	Caucasian	English	No (N/A)	None	Mostly A's
2	California	11	Female	No	Hispanic	English	No (N/A)	None	Mostly A's
3	California	11	Female	No	Other	English	No (N/A)	None	Mostly A's
4	California	11	Female	No	Caucasian	English	No (N/A)	AP Chemistry	Mostly A's
5	California	11	Female	Yes	Hispanic	English	No (N/A)	IB Honors Science	Mostly A's
6	California	11	Female	No	Hispanic	English	No (N/A)	None	Mostly B's
7	California	11	Female	No	Caucasian	English	Yes (ADHD)	None	Mostly A's
8	California	11	Male	No	Asian	English	No (N/A)	IB Biology, Chemistry	Mostly A's
9	California	11	Male	Yes	Hispanic	English	No (N/A)	None	Mostly B's
10	California	11	Female	No	Caucasian	English	No (N/A)	Chemistry	Mostly B's
11	California	11	Male	Yes	Prefer not to answer	English	No (N/A)	None	Mostly B's
12	California	11	Male	No	Caucasian	English	No (N/A)	None	Mostly B's
13	Connecticut	10	Female	–	African American	–	–	–	High Achieving
14	Connecticut	11	Male	–	Caucasian	–	–	–	High Achieving
15	Connecticut	12	Female	–	Hispanic	–	–	–	High Achieving

Note. *Parent report of science grades or teacher estimate of achievement level.
–: Missing data

Appendix 4-E

Braille Cognitive Lab Report

Cognitive Lab Study: Accessibility of Science Clusters for Braille Readers

Fran Stancavage

Susan Cole

April 2019

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1. INTRODUCTION

This set of cognitive labs was designed to determine if students using braille can understand the task demands of selected interactive Next Generation Science Standards (NGSS)-aligned science clusters and navigate the interactive features of these clusters in a manner that allows them to fully display their knowledge and skills relative to the constructs of interest. The clusters for the study were sampled from those that had already been selected for braille translation. The cognitive labs were designed to address the following three research questions:

1. Can students using braille provide responses to the selected interactive NGSS-aligned science clusters that are consistent with their knowledge and skills relative to the constructs of interest?
2. Within the selected clusters, can students successfully navigate all the included interaction types, or are further modifications needed to make the clusters fully accessible?
3. How much time do students using braille require to work their way through the selected clusters, and what strategies can be recommended to enable students using braille to complete clusters within a single testing session (to improve continuity)?

Although the American Institutes for Research (AIR) team was able to collect relevant data for this cognitive lab study, there were some limitations to the analysis. Most importantly, there were far fewer eligible visually-impaired students willing to participate in the study than anticipated, and some of them, although technically readers of braille, did not use braille while responding to the science questions in the cognitive labs. In addition, in several of the cognitive lab sessions, students' interactions with the clusters was hampered by technical issues with the Job Access With Speech (JAWS) screen-reading software and/or the Refreshable Braille Display (RDB) supplied locally, as well as by text-to-speech (TTS) tagging or braille embossing problems that arose in the beta-version materials. The latter were used in the cognitive labs due to the timing of the study.

2. METHODS

2.1 STUDY DESIGN

Two science clusters were sampled for each grade band (i.e., elementary, middle, and high school), and tailored protocols were developed for each cluster. The original design called for a minimum of six cognitive labs at each grade level, but due to recruitment challenges (discussed further in this section), labs were only conducted with ten students in total. The cognitive labs were held in Oregon and West Virginia between October 2018 and January 2019. The interviews lasted two hours, and each student was presented with one or both clusters for their grade band, depending on how much time the student took to complete the first cluster.

As part of the cognitive lab introductory activities, students were trained in the concurrent think-aloud technique. Using an elementary-level science cluster, which was not one of the clusters evaluated in the study, the interviewer first modeled the technique in Part A (first scored question) and then had the student practice in Part B (second scored question).

Students then moved on to their first assigned cluster. They were encouraged to think out loud as they worked through the cluster, and interviewers were instructed to use follow-up probes to clarify and expand on what the student said (or what the student was observed doing). Probes, which were tailored to the specifics of the cluster, focused on whether the student was able to find all the information needed to respond to the questions, what the student thought about the ways in which they had to enter answers to questions (for questions with innovative response formats), and if they would change anything about the way the information was presented to make it easier to work on the questions. A final probe allowed the student to report on anything else they found notable about the questions or introductory material in the cluster.

Students who were able to complete the first cluster by the 1.5-hour mark (out of the scheduled 2-hour lab) were moved on to the second cluster for their grade band. Probes were only administered after the student had completed all the questions in a given cluster in order to ensure that probing on the earlier questions would not influence the student's interactions with the later questions.[1]

Interviewers brought embossed braille forms to the cognitive labs. The site was responsible for providing other resources, such as JAWS and an RBD. AIR requested that a teacher of the visually impaired (TVI) or a teacher assistant be present in the room during the cognitive lab and assist the student as they would during an actual test. In most cases, prior to the interview, the interviewer briefly discussed with the TVI/teacher assistant what resources the student used to navigate online tests and how frequently/in what ways the TVI/teacher assistant typically assisted the student during testing. This information helped the interviewer to further tailor their probes and observations.

2.2 Interviewer Training

The project leads provided a 4-hour training for the interviewers who would be conducting the cognitive labs. Because all the interviewers were experienced in the cognitive interview technique, the training primarily focused on reviewing the content of the clusters and familiarizing the interviewers with the test platform and the specifics of the cognitive lab protocols. An assessment program manager was present at the training to provide an overview of the test platform and to respond to any technical questions.

2.3 Study Sample

Permission to recruit students for the study was secured from four states. In each state, the project manager and project director worked with relevant school and district personnel to recruit eligible students and coordinate logistics. Ultimately, only two states, Oregon and West Virginia, were able to provide students for the study.

The recruitment materials specified a need for students in grades 6, 7, 9, 10, or 12 who use braille, and all the recruited students were in fact able to use braille to some degree; however, an unanticipated complication was that some of the students who were partially sighted chose to use other resources (e.g., the Zoom tool) to navigate the clusters. Given that there were so few students

[1]To stay within the agreed-upon 2-hour time limit, the interviewer sometimes stopped the student before they finished the second cluster in order to leave sufficient time for probing.

available, the AIR team took whomever was recruited. The characteristics of the sample, by student, are shown in Table 1 below.

Students in grades 6 and 7 were administered the elementary-school-level clusters, students in grades 9 and 10 were administered the middle-school-level clusters, and students in grade 12 were administered the high-school-level clusters.

Table 1. Characteristics of Sample, by Student

Student	Grade	Gender	Resources Used in the Cognitive Lab
1	6	Male	JAWS, RBD, braille*
2	6	Female	Zoom, larger cursor
3	9	Male	Zoom, larger cursor, JAWS, braille
4	9	Male	Zoom
5	9	Male	JAWS, RBD
6	10	Male	JAWS, RBD, braille
7	10	Female	Braille, ChromeVox**
8	10	Female	Zoom
9	12	Female	Zoom, JAWS, braille
10	12	Male	Inverse colors, zoom

Note. **Braille* refers to the embossed braille forms
**ChromeVox is an alternative TTS reader.

3. Findings and Recommendations

3.1 Resources Used

The students used the available resources in a variety of ways during the cognitive labs. It was common for the students to switch between resources (e.g., moving between embossed braille, JAWS [sometimes coupled with an RBD], the Zoom tool [where relevant]). Some of the partially-sighted students chose to use only zoom, citing reasons such as having only “beginner” level braille skills or feeling that navigation using braille took longer; others switched between the Zoom tool and other resources. One TVI reported that the partially-sighted student they were assisting switched based on “eye fatigue and lighting conditions.” At least two students used the embossed braille forms almost exclusively to read the questions and reference the introductory materials, but switched to JAWS to enter their answers. One of these students reported that they used the embossed braille forms because it was easier than scrolling up and down the page using JAWS. Another partially-sighted student used the embossed braille forms and a screen reader similar to JAWS, but they also looked very closely at the screen to see where to place the cursor when responding to the questions.

Two students, one assigned to a middle school cluster and the other assigned to a high school cluster, reported that they would normally be offered a Perkins Brailler (also called Perkins Braille Writer) to take notes during testing. The AIR team did not anticipate or provide this resource, which is the equivalent to scratch paper for a braille user and is a standard accommodation for visually-impaired students in testing situations. It can also be used by the student to type the answers in braille, after which the TVI/teacher assistant can transcribe the answers and enter them into the test system.

3.1.1 Hardware and Software Resources

As mentioned previously, there were technical issues with some of the locally-supplied resources used in the cognitive labs. In both states, JAWS often did not work smoothly, and there were instances in which the RBD did not operate at all. As a result, some of the students struggled more with navigation than they usually would. In a couple of cases, these students reported depending more on the TVI/teacher assistant and embossed braille forms than they normally would have.

One TVI noted that every difficulty that their student encountered had come up in a real testing situation—problems with the RBD crashing, unpredictable behavior with JAWS, and “bad” embossed braille forms. The TVI said that, even when everything is tested in advance (as the RBD is), resources still do not necessarily work inside AIR’s test delivery system (TDS).

3.1.2 Embossed Braille Forms

Students were generally taken aback when they first realized the number of pages in the embossed braille forms, and, with no prior exposure to the science clusters, they had not anticipated or prepared for the need to keep track of information across multiple pages. Most of the other challenges that students experienced with this resource arose from inadvertent errors in the beta-version forms. Some of these errors were fixed after the first cognitive lab, but others persisted. In a normal cognitive lab study with a larger subject pool, all protocols would be pilot tested, which would have offered an opportunity to fix problems like this before the materials were used in the actual study.

However, some students also reported encountering graphical elements that—as rendered—were difficult to discriminate on the embossed forms. For example, one student reported that it was hard to differentiate between the two graph lines that, in the print version, were distinguished by different tones of grey. Another student indicated that it was difficult to discern the overall layout of a map of the United States, in which some states were highlighted for sharing a common characteristic, because the state lines, the line marking the boundary of the United States, and the lines outlining the Great Lakes were all too similar.

Regardless of these various issues, most students felt that the braille forms were easier to work with than using JAWS.

3.1.3 JAWS and Other Online Navigation Issues

There were significant problems with JAWS that prolonged the time it took students to work through the clusters. Some of these problems were caused by TTS-formatting configuration errors that were not caught in advance, but others had to do with the way in which JAWS was set up by the TVI/teacher assistant. An example of the latter was an instance in which JAWS was accidentally set to read all the navigation marks and not just the substance of the text. Proper settings are covered in the *Braille Requirements and Testing Manual*, but were not discussed with the TVIs/teacher assistants who were preparing for the cognitive labs.

Other challenges were caused by conventions with which the students were not familiar. In particular, students often appeared confused when JAWS skipped over a table or figure that had been judged as too complex to be read successfully by JAWS. It might have been helpful if the TTS tagging had included embedded text that instructed students to switch to the screen image or the embossed braille forms in order to see the contents of the table or figure.

For tables that were read by JAWS, at least one student noted that it would be helpful for JAWS to indicate when the table was entered and exited, rather than just reading "table of checkboxes" multiple times as it progressed through the table; however, it was not clear whether the student had JAWS set up correctly.

Several students had difficulties using the Tab key effectively, repeatedly finding themselves in some other location than they expected when they tabbed forward or back. There seemed to be some interaction between problems with tabbing and the students' confusion about JAWS not reading the tables and figures (however, it should be noted that one student, who did not have any problems navigating with JAWS, said that it would have been very helpful to be able to easily tab between the question stem and the response fields so that students could quickly review the question—potentially multiple times—as they considered their response).

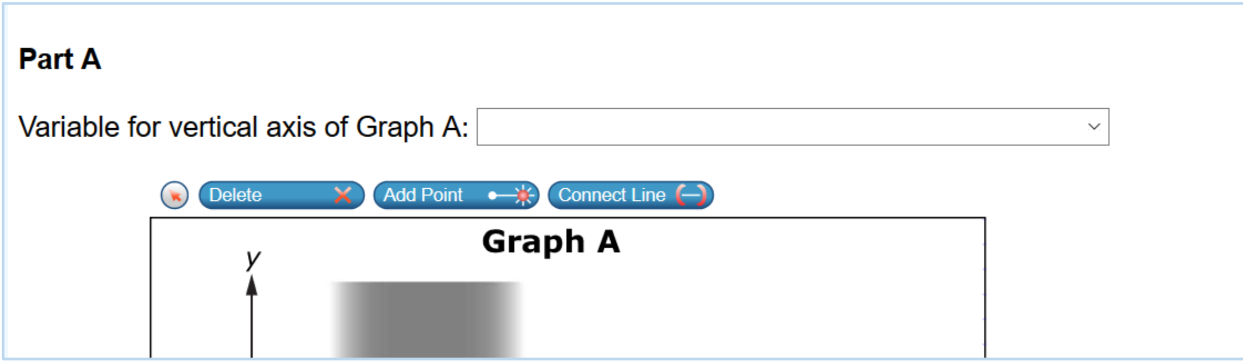
Finally, there were issues associated with the way in which drop-down boxes were handled by JAWS. Some students were not familiar with the term "combo boxes," which was used to describe these boxes, and many students were confused by the ways in which JAWS handled the response options for these boxes. In some cases, it appeared that JAWS did not read these choices at all (which was consistent with the current TXX business rules), while in other cases JAWS read the options, but only after a response was selected. Finally, the tagging may have been inadequate, as at least one student didn't understand what JAWS was reading until the TVI showed them where the various parts of the question were, especially the text in the drop-down boxes.

3.1.4 Zoom Tool

Students who used the Zoom tool did not encounter many problems applying this tool to the science clusters, although one student failed to discern at least one drop down box as they moved through the text. These students did, however, suggest several modifications that they felt would improve their experience, including the following:

- Enable the user to change the size of tables or images on all sides rather than just two sides to avoid having to scroll sideways.
- Add additional spacing in the text; at x3 or greater zoom, the spacing is too tight.
- Make the sizing of the answer buttons consistent when zoomed in—currently the answer buttons on the multiple-choice questions stayed small, whereas other answer buttons got larger when zoomed in.
- To help with viewing the drop-down boxes (see example in Figure 1), format the boxes with high contrast or a thicker line.

Figure 1. Example Drop-Down Box



3.1.5 Assistance from the TVI/Teacher Assistant

The level of TVI/teacher assistance varied in relation to the student's fluency with the other resources. An added factor in the level of assistance provided to students in the cognitive labs was the failure of the RBDs in some sessions. Without the RBD, students who could not see the computer screen required assistance to enter their responses.

The most facile student in our sample, who was very comfortable using both the embossed braille forms and JAWS, still asked for some assistance from the TVI, particularly with online navigation. At the other end of the scale, the following vignette illustrates how one TVI worked with a student who needed considerable support.

Example of a TVI assisting a student who was not very facile with the other resources available.

One student began by letting JAWS read through the entire introduction and most of the questions before asking if they could pause it. The TVI gave the student the instructions to do so. The student said that they were being hit with too much information at once, so they asked for the embossed braille form. The TVI found the first page and directed the student through most of the content, reading a lot of it out loud. The TVI noted that this was an official accommodation that the student was allowed to use during tests. The student had difficulty reading the braille out loud–stumbling over words and parts of words and asked the TVI for a lot of help with the figures. When the student had trouble reading Table 1 (included in the introduction) on the braille form, they decided to go back to JAWS. JAWS jumped ahead to Table 2 (part of the first scorable question), and it took some effort for the student to go back to Table 1. The TVI helped the student find Table 1, and the student followed along on the braille form as JAWS read the text preceding Table 1 out loud; however, JAWS did not read Table 1, instead skipping to the next paragraph of text. The student wanted to try typing on the keyboard to see if it would help bring up the table, but the TVI explained that there was no text box to type anything into. The TVI suggested that the student tab forward. The TVI said that in a real test situation, she would offer to read the table at this point. The student said this would be helpful, and the interviewer indicated that this was acceptable, so the TVI read the table out loud while the student followed along on the braille form.

3.2 General Accessibility Issues

An accessibility issue that, although it primarily affects the embossed braille forms, also has implications for screen layout, has to do with the inconsistent locations in which cluster components (e.g., questions, tables and figures, other text) appear on the page. Without the ability to quickly discern the overall layout of each page or screen, it was much harder for students in the study to process the information being conveyed. One student mentioned that it would be helpful if question stems consistently appeared on the top of the page, as in some cases the display that follows the item identifier (e.g., Part A) starts with a table or other graphic, with the text of the item stem following. Given the student feedback, it would be better to position the table/graphic below the item stem. Another student was observed to completely overlook a short paragraph of text that appeared between two large graphics in the introduction. Moreover, there were no sufficient cues to alert the student to the fact that they had missed an element. When blocks are being prepared for braille readers and other visually impaired students, it would be helpful to take these considerations into account and modify the page and screen layouts accordingly.

Similarly, one student's thoughts about how they would use the various resources to efficiently work through the science clusters (see graphic below), suggest another modification that would help maximize accessibility.

Thoughts from a student on how to best use resources to work through the science clusters.

Both the student and their TVI noted that working with the embossed braille forms for the science clusters was a departure from their usual testing experience because most traditional test questions can be rendered on a single page. Upon reflection, the student said that the strategy that would work best for them would be to

- first read through the whole cluster using the embossed braille form; and then
- navigate the questions with JAWS and an RBD, referring back to text passages as needed using these tools; however, where there was a need to refer back to a figure or chart, use the embossed braille.

The student indicated that to successfully carry out this strategy, they would need a better system for keeping all the braille pages organized so as to be able to quickly access the necessary graphics. Providing an index, or some form of page headers, might help with this problem.

3.3 Timing and Continuity

One of the goals at the beginning of the study was to determine whether students could complete an entire cluster during a single testing session; the results suggest that timing will not be a major issue, so long as schools are able to provide uninterrupted 1-hour testing sessions, if necessary. Despite the technical issues with JAWS, the RBD, and the braille forms, all but two of the students were able to complete at least one of the clusters during the cognitive labs, and one of the students who failed to complete the cluster was not focused or motivated to respond to the questions. The labs were approximately 1.5 hours long, not including the introduction and think-aloud modeling and practice. Given that they involved thinking aloud and probing, as well as working the questions, 1-hour testing sessions should be sufficient for actual administrations.

4. Conclusions

In general, both the students who relied entirely on braille and/or JAWS and those who had some vision and were able to read the screen with the Zoom tool were able to find the information they needed to respond to the questions, navigate the various response formats, and finish within a reasonable amount of time. To varying degrees, assistance from the TVI/teacher assistant was necessary, but this was most likely not qualitatively different from the assistance that would be provided on a more traditional test.

However, the clusters were clearly different from (and more complex than) other tests with which the students were familiar, and students should be given adequate time to practice with at least one sample cluster before taking the state test. It would also be helpful for students to work with their TVIs/teacher assistants in advance to develop a strategy for organizing and using the information required to answer the test questions. For example, students might want to take notes on a Perkins Brailler as they work. Given that the challenges of the science clusters are not unlike the challenges that students are likely to encounter under curricula based on NGSS or Common Core State Standards (CCSS) or their equivalent, students could be expected to become more fluent in the requisite skills as such curricula become more widespread.

Because of the large numbers of substantively important figures and tables in the clusters, we judge the embossed braille forms to be essential for any student who cannot see the material on the screen with magnification. Embossing is already set to “automatic” on all AIR science tests; however, in the case of the science clusters, test administrators (TAs) should be instructed to have the forms available before the student begins work on a given cluster, as the embossing would otherwise be very disruptive.

A major challenge that we observed in the cognitive labs—which would apply to more conventional tests, as well—was the temperamental functioning of JAWS and the RBDs. There were multiple instances of these resources failing during the cognitive labs, even when they had been tested in advance. This might be avoided with more rigorous user acceptance testing (UAT) of items using JAWS, but it also might require changes at the local level, such as better training for TVIs/teacher assistants or better maintenance of the devices.

Among the innovative response formats encountered in the science clusters that were used in the cognitive labs, the drop-down boxes proved to be the most problematic (specifically for students who were trying to navigate the science clusters using JAWS), since the drop-down options were not tagged to be read by JAWS. AIR should consider changes to the business rules in order to allow the drop-down options to be read.

The following recaps the tool-specific recommendations offered in the report.

For braille forms,

- make sure that graphic elements, such as graph or map lines, are bold enough or sufficiently contrasted to be easily discriminated;

- consider reformatting so that page layout is more predictable (e.g., always keeping text together rather than interspersing it with large graphics); and/or
- consider adding an index or page headers to make it easier for students to keep track of information across multiple sheets of embossed braille.

For JAWS,

- provide more cues when a student needs to switch to the braille form or the screen image to view a table or figure that JAWS will skip over;
- add navigation markers to indicate when the reader is entering or exiting a table if tables are tagged to be read by JAWS; and/or
- provide a way for the student to readily tab between the question stem and the response field(s).

For the Zoom tool,

- enable the user to change the size of tables or images on all sides rather than just two sides to avoid having to scroll sideways;
- add additional spacing in the text; at x3 or greater zoom, the spacing is too tight;
- make the sizing of the answer button consistent when zoomed in—as currently configured, the answer buttons on the multiple-choice questions stay small, whereas other buttons get larger when zoomed in; and/or
- format the boxes with high contrast to help with viewing the drop-down boxes.

Montana Science Assessment

2022–2023

Volume 5:
Test Administration

OFFICE OF PUBLIC INSTRUCTION
ELSIE ARNTZEN, STATE SUPERINTENDENT

TABLE OF CONTENTS

LIST OF TABLES

LIST OF APPENDICES

1. Test Intervals, Options, and Administrative Roles

1.1 Testing Windows

The 2022–2023 Montana Science Assessment (MSA) testing window spanned twelve weeks for the summative assessments. The paper-pencil fixed-form tests for summative assessments were administered concurrently during the online summative windows.

1.2 Test Options and Administrative Roles

The MSA is administered primarily online. To ensure that all eligible students in the tested grades were given the opportunity to take the MSA, a number of assessment options were available for the 2022–2023 test administration to accommodate students' needs. Table 1 lists the testing options that were offered in 2022–2023. Once a testing option was selected, it applied to all tests in the content area.

Table 1. Summary of Tests and Testing Options in 2022–2023

Assessments	Testing Options	Test Mode
Summative	English	Online
	Spanish (Toggle)	Online
	Braille	Online
	Paper-Pencil Braille Fixed-Form Test[a]	Paper-Pencil

[a]For the paper-pencil fixed-form tests, all student responses were entered into the Data Entry Interface (DEI) by test administrators.

To ensure standardized administration conditions, Test Administrators (TAs) followed procedures outlined in the *MSA Test Administration Manual* (TAM). TAs reviewed the TAM prior to the beginning of testing in order to ensure that the testing room was prepared appropriately (e.g., removing certain classroom posters, arranging desks). Make-up procedures were established for any students who were absent on the day(s) of testing. TAs followed required administration procedures and directions and read the boxed directions verbatim to students, ensuring standardized administration conditions.

1.2.1 Administrative Roles

The key personnel involved with the test administration for the Montana Office of Public Instruction (OPI) are Authorized Representatives (ARs), System Test Coordinators (STCs), Building Coordinators (BCs), Test Administrators (TAs), and Proctors (PRs). The main responsibilities of these key personnel are described in the following subsections. More detailed descriptions can be found in the TAM provided online at https://mt.portal.cambiumast.com/resources/.

Authorized Representative

The AR may add users with STC roles in the Test Information Distribution Engine (TIDE). For example, a director of special education may need STC privileges in TIDE to access district-level data for the purposes of verifying test settings for designated supports and accommodations. ARs have the same test administration responsibilities as STCs. Their primary responsibility is to coordinate the administration of the MSA in the district.

System Test Coordinator

The STC is primarily responsible for coordinating the administration of the MSA at the district level.

STCs are responsible for the following:

- Reviewing all MSA policies and test administration documents
- Reviewing scheduling and test requirements with BCs and TAs
- Working with BCs and Technology Coordinators (TCs) to ensure that all systems, including Cambium Assessment, Inc.'s Secure Browser, are properly installed and functional
- Importing users (including BCs and TAs) into TIDE
- Verifying all student information and eligibility in TIDE
- Scheduling and administering training sessions for all BCs, TAs, and TCs
- Ensuring that all personnel are trained on how to properly administer the MSA
- Monitoring the secure administration of the tests
- Investigating and reporting all testing improprieties, irregularities, and breaches reported by the TAs
- Attending to any secure material according to OPI policies

Building Coordinator

The BC is primarily responsible for coordinating the administration of the MSA at the school level and ensuring that testing within his or her school is conducted in accordance with the test procedures and security policies established by the OPI.

BC responsibilities include the following:

- Based on test administration windows, establishing a testing schedule with STCs and TAs
- Working with technology staff to ensure timely computer setup and installation

- Working with TAs to review student information in TIDE to ensure that student information and test settings for designated supports and accommodations are correctly applied
- Identifying students who may require designated supports and test accommodations and ensuring that procedures for testing these students follow OPI policies
- Attending all district trainings and reviewing all OPI policies and test administration documents
- Ensuring that all TAs attend school or district trainings and review online training modules posted on the portal
- Establishing secure and separate testing rooms if needed
- Downloading and planning the administration of the classroom activity with TAs
- Monitoring the secure administration of the tests
- Monitoring testing progress during the testing window and ensuring that all students participate, as appropriate
- Investigating and reporting all testing improprieties, irregularities, and breaches reported by the TAs
- Attending to any secure material according to OPI policies

Test Administrator

A TA is primarily responsible for administering the MSA. A TA will have access to their own students' results when they are made available in the reporting System.

TAs are responsible for the following:

- Completing MSA test administration training
- Reviewing all OPI policy and test administration documents before administering the MSA
- Viewing student information before testing to ensure that a student receives the proper test with the appropriate supports and reporting any potential data errors to BCs and STCs, as appropriate
- Administering the MSA
- Reporting all potential test security incidents to the BCs and STCs in a manner consistent with OPI and district policies

1.2.2 Standard Online Test Administration

Within Montana's testing window, schools can set testing schedules, allowing students to test in intervals (e.g., multiple sessions) rather than in one long test period, which minimizes the interruption of classroom instruction and efficiently uses the test's facility. With online testing, schools do not need to handle test booklets and address the storage and security problems inherent in large shipments of materials to a school site.

BCs oversee all aspects of testing at their schools and serve as the main point of contact, while TAs administer the online assessments only. TAs are trained in the online testing requirements and the mechanics of starting, pausing, and ending a test session. Training materials for the test administration are provided online.

To start a test session, the TA must first enter the TA Interface of the online testing system using his or her own computer. A session ID is generated when the test session is created. Students who are taking the assessment with the TA must enter their State Student Identifier (SSID), first name, and session ID into the Student Interface using computers provided by the school. The TA then verifies that the students are taking the appropriate assessments with the appropriate accessibility feature(s). Students can begin testing only after the TA has confirmed the settings. The TA then reads aloud the directions in *Section IX. Day of Test Administration* of the TAM to the students and guides them through the login process.

Once an assessment has started, the student must answer the test question presented on a page before proceeding to the next page. Skipping items is not permitted. For the online computer test, students are allowed to scroll back to review and edit previously answered items, if these items are in the same test session and the session has not been paused for more than 20 minutes. Students may review and edit responses they have previously provided before submitting the assessment. During an active online computer test session, if a student reviews and changes the response to a previously answered item, then all items that follow to which the student already responded remain the same. If a student changes the answers, no new items are assigned. For example, a student pauses for 10 minutes after completing Item 10. After the pause, the student goes back to Item 5 and changes the answer. If the response change in Item 5 changes the item score from wrong to right, the student's overall score will improve; however, there will be no change in Items 6–10.

For the summative test, an assessment can be started on one day and completed on another. For the online computer test, the assessment must be completed within 45 calendar days of the start date, or the assessment opportunity will expire.

During a test session, TAs may pause the test for a student or group of students to take a break. It is up to the TAs to determine an appropriate stopping point; however, to ensure the integrity of test scores or testing, the online computer test cannot be paused for more than 30 minutes. If that happens, the student must begin a new test session, which starts from where the student left off. Previous responses are no longer available for viewing or editing.

The TAs must remain in the room at all times during a test session in order to monitor student testing. Once the test session ends, the TAs must ensure that each student has successfully logged out of the system. Then, the TAs must collect any handouts or scratch paper that students used during the assessment and send them for secure shredding.

1.2.3 Spanish Online Test Administration

The Spanish versions of the MSA are provided as an accommodation for students whose primary language is Spanish and who use dual language supports in the classroom. For Montana students, Spanish versions of the tests are offered online. Students with this support can toggle between Spanish and English display of the item by clicking a button at the top of the screen inside the test.

1.2.4 Braille Online Test Administration

The fixed-form braille versions of the MSA are available with the same test blueprint as the English and Spanish versions of the test.

The braille interface is as follows:

- The braille interface includes a text-to-speech component for mathematics consistent with the read-aloud assessment accommodation. The Job Access With Speech (JAWS) screen-reading software provided by Freedom Scientific is an essential component that students use with the braille interface.
- Science items are presented to students in Unified English Braille (UEB) Contracted with Nemeth Braille code.

Before administering the online summative assessments using the braille interface, TAs must ensure that the technical requirements are met. These requirements apply to the student's computer, the TA's computer, and any supporting braille technologies used in conjunction with the braille interface.

1.2.5 Braille Paper-Pencil Test Administration

The Braille paper-pencil versions of the MSA are provided as an accommodation for students who are visually impaired.

The AR must order the accommodated test materials on behalf of the students who need to take the Braille paper-pencil test with the OPI. Once approved, the testing contractor is notified and ships the appropriate test booklets and the Braille paper-pencil TAM to the district.

After the student has completed the assessment, the TAs enter the student responses into the DEI and return the test booklets to the testing vendor. The tests submitted via the DEI are then scored.

2. Training and Information for Test Coordinators and Administrators

Authorized Representatives (ARs), System Test Coordinator (STCs), and Building Coordinators (BCs) oversee all aspects of testing at their schools and serve as the main points of contact, while Test Administrators (TAs) administer the online assessments. The online Cambium Assessment, Inc. (CAI) TA Certification Course, webinars, user guides, manuals, and training sites are used to train TAs in the online testing requirements and the mechanics of starting, pausing, and ending a test session. Training materials for test administration are provided online.

2.1 Online Training

Multiple online training opportunities are offered to key staff.

2.1.1 TA Certification Course

CAI's online TA Certification Course is available to any user in the Test Information Distribution Engine (TIDE). This web-based course is about 30–45 minutes long and covers information on testing policies and steps for administering a test session in the online system. This interactive course requires participants to start test sessions under different scenarios. Throughout the training and at the end of the course, participants are required to answer multiple-choice questions about the information provided.

2.1.2 Practice Site

In December 2019, a practice site was opened for TAs and students. TAs could practice administering assessments and starting and ending test sessions on the TA Training Site, and students could practice taking a short online assessment on the Student Practice and Training Site. A student could log in directly to the Student Practice and Training Site as a guest without a TA-generated test session ID, or the student could log in through a practice test session created by the TA.

The MSA practice tests contain the same item types (stand-alone and clusters) students would encounter on the MSA. The practice tests are designed to provide students and teachers with opportunities to quickly familiarize themselves with the software and navigational tools they would use for the MSA. Practice tests were organized by grade bands (grades 5 and 8). The practice tests were refreshed in August 2021.

2.1.3 Manuals and User Guides

The *Test Administration Manual* (Appendix 5-A) provides information for TAs administering the Montana Science online summative assessments, including policies and procedures. It includes screen captures and step-by-step instructions on how to administer the online tests.

The *Assistive Technology Manual* (Appendix 5-B) provides an overview of the embedded and non-embedded assistive technology tools that can be used to help students with specific accessibility needs complete online tests in the Test Delivery System (TDS). It includes lists of supported devices and applications for each type of assistive technology that students may need, as well as setup instructions for the assistive technologies that require additional configuration in order to work with the TDS.

The *Technology Guide* on the State Portal[1] outlines Technology required for administering an online assessment, including operating system requirements. This guide also provides detailed

[1] https://mt.portal.cambiumast.com/tech-guide.html

instructions for configuring and troubleshooting the Secure Browser on supported operating systems used for online assessments, and information on configuring workstations and networks.

The *Test Information Distribution Engine User Guide* (Appendix 5-C) was designed to help users navigate TIDE. It provides information on managing user account information; student account information; student test settings and accommodations; appeals; and voice packs.

The *Reporting System User Guide* (Appendix 5-D) provides information about the Reporting System, including instructions for viewing score reports, accessing test management resources, creating and editing rosters, and searching for students.

The *Test Administrator User Guide* (Appendix 5-E) was designed to help users navigate the TDS, including the Student Interface and the TA Interface, and to help TAs manage and administer online testing for students.

All manuals and user guides pertaining to the 2022–2023 online assessments were available on the Montana portal at: https://mt.portal.cambiumast.com/. ARs, STCs, and BCs used the manuals and user guides to train TAs in test administration policies and procedures.

2.2 System Test Coordinator Training Workshops

System Test Coordinator (STC) training webinars were held monthly for the 2022–2023 school year and during a 3-day workshop in October 2022. During the trainings, STCs were provided with information to support training of the TAs.

3. Test Security

All test items, test materials, and student-level testing information are considered secure materials for all assessments. The importance of maintaining test security and the integrity of test items is stressed throughout the webinar trainings and in the user guides, modules, and manuals. Features in the testing system also protect test security. This section describes system security, student confidentiality, and policies on testing improprieties.

3.1 Student-Level Testing Confidentiality

All secure websites and software systems enforce role-based security models that protect individual privacy and confidentiality in a manner consistent with the Family Educational Rights and Privacy Act (FERPA) and other federal laws. Secure transmission and password-protected access are basic features of the current system and ensure authorized data access. All aspects of the system—including item development and review, test delivery, and score reporting—are secured by password-protected logins. Our systems use role-based security models to ensure that users can access only the data to which they are entitled and can edit data only in accordance with their user rights.

The following are the three dimensions related to ensuring that students are accessing appropriate test content:

1. *Test eligibility* refers to the assignment of a test to a particular student.

2. *Test accommodation* refers to the assignment of a test setting to specific students based on needs.
3. *Test session* refers to the authentication process of a Test Administrator (TA) creating and managing a test session, the TA reviewing and approving a test (and its settings) for every student, and the student signing on to take the test.

FERPA prohibits public disclosure of student information or test results. The following are examples of prohibited practices:

- Providing login information (username and password) to other authorized Test Information Distribution Engine (TIDE) users or to unauthorized individuals
- Sending a student's name and Statewide Student Identifier (SSID) together in an email message; if information must be sent via email or fax, include only the SSID, not the student's name
- Having students log in and test under another student's SSID

Test materials and score reports should not be exposed to identify student names with test scores except by authorized individuals with an appropriate need to know.

All students, including homeschooled students, must be enrolled, or registered at their testing schools to take the online, paper-pencil, or braille assessments. Student enrollment information, including demographic data, is generated using a Montana Office of Public Instruction (OPI) file and uploaded nightly via a secured file transfer site to the Reporting System during the testing period.

Students log in to the online assessment using their legal first name, SSID, and a test session ID. Only students can log in to an online test session; TAs, proctors, or other personnel are not permitted to log in to the system on behalf of students, although they are permitted to assist students who need help logging in. For the paper-pencil versions of the assessments, TAs are required to affix the student label to the student's answer document.

After a test session, only staff with the administrative roles of Authorized Representative (AR), System Test Coordinator (STC), Building Coordinator (BC), or the Teacher can view their students' scores. TAs do not have access to student scores unless they are the student's assigned teacher.

3.2 System Security

The objective of system security is to ensure that all data are protected and accessed appropriately by the designated user groups. System security focuses on protecting data and maintaining data and system integrity as intended, including ensuring that all personal information is secured, that transferred data (whether sent or received) is not altered in any way, that the data source is known, and that only a specific, designated user can perform any service.

A Hierarchy of Control: As described in Section 1.2.1, Administrative Roles, all ARs, STCs, BCs andTAs have defined roles and levels of access to the testing system. When the TIDE testing window opens, the OPI provides a verified list of ARs to the testing contractor, who uploads the information into TIDE. ARs are then responsible for selecting and entering the STCs' and BCs' information into TIDE, and the BC is responsible for entering TA information into TIDE. Throughout the year, the AR, STC, and BC are also expected to delete information in TIDE for any staff members who have transferred to other schools, resigned, or no longer serve as TA.

Password Protection: All access points by different roles—at the state, district, school principal, and school staff levels—require a password to log in to the system. Newly added BCs and TAs receive separate passwords assigned by the school through their personal email addresses.

Secure Browser: A key role of the Technology Coordinator (TC) is to ensure that the Secure Browser is properly installed on the computers used for the administration of the online assessments. Developed by the testing contractor, the Secure Browser prevents students from accessing other computers or Internet applications and from copying test information. The assessments can be accessed only through the Secure Browser and not by other Internet browsers. The Secure Browser suppresses access to commonly used browsers, such as Internet Explorer and Firefox, and prevents students from searching for answers on the Internet or communicating with other students.

3.3 Security of the Testing Environment

3.3.1 Duties of Testing Personnel

The BCs and TAs work together to determine appropriate testing schedules based on the number of computers available, the number of students in each tested grade, and the average amount of time needed to complete each assessment.

Testing personnel are reminded in the online training and user manuals that assessments should be administered in testing rooms that do not crowd students. Good lighting, ventilation, and freedom from noise and interruption are important factors to consider when selecting testing rooms.

TAs must establish procedures to maintain a quiet environment during each test session and recognize that some students may finish more quickly than others. If students are allowed to leave the testing room when they finish, TAs are required to explain the procedures for leaving and where students are expected to report once they leave without disrupting others. If students are expected to remain in the testing room until the end of the session, TAs are encouraged to prepare some quiet work for students to do after they finish the assessment.

If a student needs to leave the room for a brief time during testing, the TAs are required to pause the student's assessment. For the online computer test, if the pause lasts longer than 20 minutes, the student can continue with the rest of the assessment in a new test session, but the system will not allow the student to return to the items answered before the pause. This measure is implemented to prevent students from using the time outside of the testing room to look up answers.

3.3.2 Room Preparation

The room should be prepared prior to the start of the test session. Any information displayed on bulletin boards, chalkboards, or charts that students might use to help answer test questions should be removed or covered. This rule applies to rubrics, vocabulary charts, student work, posters, graphs, content-area strategies charts, and other materials. The testing personnel's and students' cell phones must be turned off and stored in the testing room out of sight. TAs are encouraged to minimize access to the testing rooms by posting signs in halls and entrances in order to promote optimum testing conditions; they should also post "TESTING—DO NOT DISTURB" signs on the doors of testing rooms.

3.3.3 Seating Arrangements

TAs should provide adequate space between students' seats. Students should be seated so that they will not be tempted to look at the answers of others. Because the online test is adaptive, it is unlikely that students will see the same test questions as other students. However, through appropriate seating arrangements, students should be discouraged from communicating with each other.

3.3.4 After the Test

At the end of the test session, TAs must walk through the classroom to pick up any scratch paper that students used and any papers that display students' SSIDs and names together. These materials should be securely shredded or stored in a locked area immediately following a test session.

For the paper-pencil versions, specific instructions on how to package and secure the test booklets to be returned to the testing contractor's office are provided in the *Test Administration Manual* (TAM).

3.4 Test Security Violations

Everyone who administers or proctors the assessments is responsible for understanding the test administration security procedures. Prohibited practices as detailed in the TAM are categorized into the following three groups:

1. *Impropriety:* This is a test security incident that has a low impact on the individual or group of students who are testing and has a low risk of potentially affecting student performance on the test, test security, or test validity (e.g., students leaving the testing room without authorization).

2. *Irregularity:* This is a test security incident that impacts an individual or group of students who are testing and may potentially affect student performance on the test, test security, or test validity. These circumstances can be contained at the local level (e.g., disruption during the test session, such as a fire drill).

3. *Breach:* This is a test security incident that poses a threat to the validity of the test. Breaches require immediate attention and escalation to the OPI. Examples include situations such as exposure of secure materials or a repeatable security/system risk. These circumstances have external implications (e.g., administrators modifying student answers, students sharing test items through social media).

District and school personnel are required to document all test security incidents in the test security incident log. The log serves as the document of record for all test security incidents and should be maintained at the district level and submitted to the OPI at the end of testing.

4. Student Participation

4.1 Eligibility

All students (including retained students) enrolled in grades 5 and 8 at public schools in Montana are required to participate in the MSA. Students must be tested in the enrolled grade assessment; out-of-grade-level testing is not allowed for the administration of the MSA.

4.2 Homeschooled Students

Homeschooled students are not public-school students and are not required to be administered state assessments.

4.3 Exempt Students

Students who are hospitalized or homebound due to illness should be tested unless there are medical constraints that prevent them from testing and the student has received an approved medical exemption from participating in the MSA.

5. Online Testing Features and Testing Accommodations

The *Usability, Accessibility, and Accommodations Guidelines* (UAAG*;* Appendix 5-F) are intended for school-level personnel and decision-making teams, including Individualized Education Program (IEP) and Section 504 Plan teams, as they prepare for and implement the MSA. The UAAG provide information for classroom teachers, English language development educators, special education, and instructional assistants to use in selecting and administering universal tools, designated supports, and accommodations for those students who need them. The UAAG are also intended for assessment staff and administrators who oversee the decisions that are made in instruction and assessment.

The UAAG apply to all students. They emphasize an individualized approach to the implementation of assessment practices for those students who have diverse needs and participate in large-scale content assessments. They focus on universal tools, designated supports, and accommodations for the MSA. At the same time, the UAAG support important instructional decisions about accessibility and accommodations for students who participate in the MSA.

The summative assessments contain universal tools, designated supports, and accommodations in both embedded and non-embedded versions. Embedded resources are part of the computer administration system, whereas non-embedded resources are provided outside of that system.

State-level users, System Test Coordinator (STCs), and Building Coordinators (BCs) can set embedded and non-embedded designated supports and accommodations based on their specific user role. Designated supports and accommodations must be set in the Test Information Distribution Engine (TIDE) before starting a test session.

All embedded and non-embedded universal tools will be activated for use by all students during a test session. One or more of the pre-selected universal tools can be deactivated by a Test

Administrator (TA) in the TA Interface of the testing system for a student who may be distracted by the ability to access a specific tool during a test session.

For additional information about the universal tools, designated supports, and accommodations offered in the 2022–2023 MSA test administration, refer to Appendix 5-F, *Usability, Accessibility, and Accommodations Guidelines*.

Appendix 5-A
Test Administration Manual

Montana Science Assessment

Test Administration Manual

2022–2023

Published January 30, 2023

TABLE OF CONTENTS

LIST OF TABLES

Montana Resources and Technical Support

The Montana Assessment Portal at https://mt.portal.cambiumast.com/is the home for all Montana Science Assessment administration information.

For questions regarding the online testing system or for additional assistance, please contact the Montana Help Desk at 1-877-365-7915 or https://mt.portal.cambiumast.com/contact.html.

The Help Desk is open Monday–Friday 6:00 a.m. to 6:00 p.m. Mountain Time. During these hours, staff will respond promptly to calls.

The Help Desk may be contacted for situations and questions including the following:

- Help is needed in preparing for online testing, such as downloading the Secure Browser and installing voice packs.
- The online testing environment is down or unavailable.
- User accounts are unavailable or users are unable to administer tests.
- Student information or test session is incorrect or missing.
- Issues are present with loading student data or student settings into the Test Information Distribution Engine (TIDE).

When contacting the Help Desk, provide the representative with as much detail as possible about the issue(s) encountered and the system on which it occurred. This should include the following:

- Type of device being used for the test
- Any error messages that appeared (code and description)
- Operating system and browser information
- Network configuration information
- Your contact information, including email address and telephone number
- Any relevant school information, grade level, and content area
- If the issue pertains to a student, provide the associated district or school for that student. SSID and additional student PII can only be communicated via phone with the help desk. **Do not leave any student identifying information such as a student name, SSID, and/or personal characteristics in a voicemail or email.**

For online Montana Science Assessment administration resources, please refer to the Montana Portal.

Section I. Overview of the Montana Science Assessment

I.I About the Montana Science Assessment

The Montana Comprehensive Assessment System (MontCAS) meets the testing requirements of the Every Student Succeeds Act (ESSA) and ensures that all children have the opportunity to obtain a high-quality education and reach proficiency on challenging state academic achievement standards and assessments. The Montana Science Assessment (MSA) is aligned to the Montana Content Standards in Science and will be administered to all students in grades 5 and 8.

The MSA is a computer-adaptive test, where items are selected on the fly based on the student's performance on previous items. Students will be presented with items from each of the three science disciplines: Life Sciences, Earth and Space Sciences, and Physical Sciences. Some content will be presented in item clusters that are designed to engage the student in a grade-appropriate, meaningful scientific activity aligned to a specific performance expectation. Each item cluster begins with a real-world phenomenon and includes two or more items that require students to demonstrate the ability to use the science practices, disciplinary core ideas, and cross-cutting concepts described by the performance expectation.

I.II Test Administration Roles and Responsibilities

Table 1: User Roles in the Online Testing System

User Role	Description
Authorized Representative (AR)	Authorized Representatives (AR) are assigned by a State User. If assigned, an AR can modify student records (including accommodations, designated supports, and interim test eligibility) as well as add System Test Coordinators (STC) or Building Coordinator (BC) in TIDE. All other responsibilities for this role should be assigned by the State.
System Test Coordinator (STC)	STCs are responsible for the following: • Adding BCs into TIDE • Ensuring that the BCs, TAs, and PRs in their districts are appropriately trained regarding the state and Montana Science Assessment administration and security policies and procedures • Reporting online test security incidents to the OPI via the secure centralized OPI MontCAS Application • Reporting non-participation for any student not completing the Montana Science Assessment via the OPI MontCAS Application • Making requests for accommodations beyond those routinely allowed via the OPI MontCAS Application • General oversight responsibilities for all administration activities in their district schools

User Role	Description
Building Coordinator (BC) ***Note:*** *A BC can be a principal, vice principal, technology coordinator, counselor, or other staff member. If possible, BCs should be people with non-instructional or limited instructional duties so that they can coordinate and monitor testing activity in the school.*	BCs are responsible for: • Reviewing the User Roles and System access document on the Montana portal at https://mt.portal.cambiumast.com/ prior to adding users to the school • Adding School Test Administrators (TA) and Proctors (PR) in TIDE. **NOTE: The PR role should be only explicitly assigned if a user should not have access to score report data in the Reporting System, otherwise all users should be given the TA role** • Identifying PRs and ensuring they are properly trained • Coordinating with PRs so they administer all assessments • Entering and/or verifying test settings for students • Creating or approving testing schedules and procedures for the school (consistent with state/territory and district policies) • Working with technology staff to ensure that necessary secure browsers are installed, and any other technical issues are resolved • Monitoring testing progress during the testing window and ensuring that all students participate, as appropriate • Addressing testing problems and reporting them to the STC, as needed • Mitigating and reporting all test security incidents in a manner consistent with Montana Science Assessment and district policies • General oversight responsibilities for all administration activities in their school and for all TAs and PRs
Test Administrator (TA) ***Note***: *The Montana Science assessments should be administered by individuals with an existing state/territory certification in education (e.g., teachers, school psychologists, principals, etc.).*	TAs are responsible for: • Completing Montana Science Assessment administration and reviewing all MontCAS policy and administration documents prior to administering any Montana Science assessments • Viewing student information prior to testing to ensure that the correct student receives the proper test with the appropriate supports. TAs should report any potential data errors to ARs and STCs as appropriate • Administering the Montana Science Assessments • Assuring the test environment meets the specific requirements • Reporting all potential test security incidents to their AR, STC, and BC in a manner consistent with state and district policies
Proctor (PR) *District or school personnel responsible for administering the Montana Science assessments in a secure manner in compliance with the policies and procedures outlined in the Online, Summative Test Administration Manual.* ***Note:*** *All roles AR/STC, BC, and TA can administer tests. Users with the role PR in TIDE will NOT be able to see score report data.*	PRs are responsible for: • completing assessment administration training and reviewing all MontCAS policy and administration documents prior to administering any Montana Science assessments • Viewing student information prior to testing to ensure that the correct student receives the proper test with the appropriate supports. TEs should report any potential data errors to ARs and STCs as appropriate • Providing students with opportunities to become familiar with the online test tools by reviewing the practice and training tests • Administering the Montana Science Assessments • Reporting all potential test security incidents to their BC, AR, and STC in a manner consistent with state and district policies

This manual refers to Test Administrators (TAs) generically whenever discussing administration or the user guide. Proctors can be defined as district or school personnel responsible for administering the Montana Science Assessments in a secure manner in compliance with the policies and procedures outlined in the *Online, Summative Test Administration Manual*. If the user role School Test Administrator in TIDE is being referenced, that will be clarified.

I.III Test Administrator Training

Prior to administering a test, STCs, BCs, and TAs (and any other individuals who will be administering any secure Montana Science Assessment) should:

- Read this manual
- Read the MSA Usability, Accessibility, and Accommodations Guidelines
- Read the Test Administrator User Guide
- View the associated training modules

The OPI expects all persons involved in handling test materials and test administration to complete the following:

- Complete the MontCAS Test Security readings and tasks as described on the OPI Test Security webpage.
- Complete the Test Security Agreement (TSA) upon activating your TIDE account. This must be signed virtually before accessing any locked tile on the portal. If a user has access to the secure tiles on the portal it signifies that they have signed their TSA.
- Complete the assessment-specific MontCAS Accessibility and Accommodation Guideline readings and tasks as described on the Usability, Accessibility and Accommodations Guidelines.
- Complete the Test Administrator (TA) Certification Course.

All Montana Science manuals and resources can be found on the Montana portal under the Resources tab (https://mt.portal.cambiumast.com/resources) and the OPI MontCAS site (http://opi.mt.gov/Leadership/Assessment-Accountability/MontCAS).

Section II. Test Administration Resources

This Test Administration Manual (TAM) for summative assessments is intended for staff who play a role in the administration of Montana Science Assessment (ARs, STCs, BCs, and TAs). This manual provides procedural and policy guidance to implement the Montana Science Assessments. The appendices of this manual contain important information that can be used as stand-alone materials and are easily extracted for printing or distribution. For a list of frequently used terms associated with the Montana Science Assessments, see Appendix A. For specific questions not addressed in this manual, please contact the Help Desk.

II.I Montana Science Assessment Administration Resources

The TAM is designed to complement a variety of other resources. All resources can be found on the Montana portal located at https://mt.portal.cambiumast.com/.

Table 2: Manuals

Resource	Description
Online Technology Guide	The *Online Technology Guide* provides staff with the technical specifications and configurations for online testing, including information on Internet, network, hardware and software requirements, and the text-to-speech function. It also provides instructions for installing and configuring the secure browser on supported operating systems, as well as information about supported operating systems and required hardware and software for using assistive technology with the secure browser. It provides configuration requirements and recommendations for frequently used assistive technologies. It also provides information on how to configure Job Access with Speech (JAWS), navigating an online test with JAWS, and how to administer a test to a student requiring Braille.
Test Administrator User Guide	The *Test Administrator User Guide* supports individuals using the test delivery system applications to manage testing for students participating in the summative assessments. This resource provides information about the test delivery system, including the Test Administrator and student applications.
TIDE User Guide	The Test Information Distribution Engine (TIDE) User Guide is designed to help users navigate the Test Information Distribution Engine (TIDE). Users can find information on managing user account information, managing student account information and student test settings and accommodations.
TIDE Account Maintenance Quick Start Guide	This Quick Start Guide provides instructions for key tasks required for proper maintenance of the TIDE User accounts.
Usability, Accessibility, and Accommodations Guidelines	This manual provides guidelines for educators in Montana to use for the administration of usability and accessibility supports for instruction and assessment of all students, including students with disabilities, English language learners (ELLs), ELLs with disabilities, and students without an identified disability or status for the following assessments: Smarter Balanced ELA and Math Assessments and Montana Science Assessments.

Resource	Description
User Roles and Access Document	This document contains a list of user roles and access for the tasks that District Users and School Users can perform in all online Montana Science Assessment systems. It is recommended to review this document prior to assigning a system user permission to individuals in your districts or schools.

Table 3: Training Modules

Module Name	Primary Audience	Objective
Student Interface Training Module	Students, District/ School Test Coordinators, Test Administrators, teachers	This module explains how to navigate the Student Interface.
Test Administrator (TA) Interface for Online Testing	District/School Test Coordinators, Test Administrators, teachers	This module presents an overview on how to navigate the Test Administrator Interface.
TIDE	District/School Test Coordinators, Test Administrators	This module provides detailed information on how to upload student information, manage users, and use other features of the TIDE system.

Table 4: Practice Test

Resource	Description
Practice Test*	Practice Tests include items and performance tasks for grades 5 and 8. Practice Tests provide a sample of the item types included in the Montana Science assessments. This is the link to the Practice Tests: https://mt.portal.cambiumast.com/students.html. See also Appendix C for additional information about the Practice Tests.

* The Practice and Training Tests can be accessed without login credentials by signing in as a guest; however, if users want to access either of these sites as a test administrator (required if they want to administer a braille Practice or Training Test), a login will be required. Contact the Help Desk if you do not have login credentials. The Practice and Training Tests do not require use of the secure browser, but it is recommended some accessibility features (such as text-to-speech) be sampled through the secure browser to verify compatibility.

Section III. Ensuring Test Security

The security of assessment instruments and the confidentiality of student information are vital to maintaining the validity, reliability, and fairness of the results.

All Montana Science Assessment items and test materials are secure and must be appropriately handled. Secure handling protects the integrity, validity, and confidentiality of assessment items, prompts, and student information. Refer to Appendix E for examples of issues that are considered violations of test security.

III.I Security of the Test Environment

Table 5 describes security requirements for the test environment during various stages of testing. The test environment refers to all aspects of the testing situation and includes what a student can see, hear, or access (including access via technology).

Individuals who have witnessed, been informed of, or suspect the possibility of a test security incident that could potentially affect the integrity of the assessments must report this incident immediately to the BC so he or she may notify the OPI via the secure Montana Comprehensive Assessment System (MontCAS) Test Irregularity Reporting System.

Table 5: Requirements of the Test Environment

Requirement	Description
BEFORE TESTING	
Instructional materials removed or covered	Instructional materials **must be removed or covered, including but not limited to** information that is displayed on bulletin boards, nametags, chalkboards or dry-erase boards, or on posters (e.g., wall charts that contain literary definitions, maps, formulas, etc.) that might assist students in answering questions.
Student seating	Students must be seated so there is enough space between them to minimize opportunities to look at each other's work, or they should be provided with table-top partitions.
Signage	If helpful, place a "TESTING—DO NOT DISTURB" sign on the door or post signs in halls and entrances rerouting hallway traffic in order to promote optimum testing conditions.
DURING TESTING	
Quiet environment	Provide a quiet environment void of talking or other distractions that might interfere with a student's ability to concentrate or might compromise the testing situation.
Student supervision	Students are actively supervised by a trained test administrator and are prohibited from access to unauthorized electronic devices that allow availability to outside information, communication among students or with other individuals outside the test environment or photographing or copying test content. This includes any device with cellular, messaging, or wireless capabilities, but is not limited to cell phones, smart watches, personal digital assistants (PDAs), iPods, cameras, and electronic translation devices.
Access to allowable resources only	Students must only have access to and use of those allowable resources identified by Montana Science that are permitted for each specific test (or portion of a test). This includes access to medical devices, which may be integrated into other technology devices. Use of such devices may require additional monitoring or a separate test setting to maintain test security.

Requirement	Description
Access to assessments	Unauthorized staff or other adults must not be in the room during testing. Only students who are testing can view items. Students who are not being tested may not be in the room where a test is being administered and must not have access to secure testing materials including test items. Based on the item type (i.e., clusters), trained test administrators may also have limited exposure to items in the course of properly administering the assessments; however, test administrators and other trained staff must not actively review or analyze any test items.
Testing through secure browser	Administration of the MSA is permitted only through the Student Interface via a secure browser or a method of securing the student device.
DURING AND AFTER TESTING	
No access to responses	ARs, STCs, BCs, TAs, PRs, and other staff are not permitted to review student responses in the testing interface or students' notes on scratch paper.
No copies of test materials	Unless needed as a print-on-demand or braille accommodation, no copies of the test items, stimuli, reading passages, or writing prompts may be made or otherwise retained.
No access to digital, electronic, or manual devices	No digital, electronic, or manual device may be used to record or retain test items, reading passages, or writing prompts. Similarly, these materials must not be discussed with or released to anyone via any media, including fax, email, social media websites, etc.
No retaining, discussing, or releasing test materials	Descriptions of test items, stimuli, printed reading passages, or writing prompts must not be retained, discussed, or released to anyone.
No reviewing, discussing, or analyzing test materials	ARs, STCs, BCs, TAs, PRs, and other staff must not review, discuss, or analyze test items, stimuli, reading passages, or writing prompts at any time, including before, during, or after testing. Students should be informed that they may not discuss or share test items, stimuli, reading passages, or writing prompts with anyone during or after testing.
All test materials must remain secure at all times	Printed materials from the print-on-demand accommodation, scratch paper, printed TIDE SSID tickets, and documents with student information must be kept in a securely locked room or locked cabinet that can be opened only with a key or keycard by staff responsible for test administration.
AFTER TESTING	
No test materials used for instructions	Test items, stimuli, reading passages, or writing prompts must not be used for instruction.
Destroy test materials securely	Printed test items/passages, including embossed braille printouts, and scratch paper must be collected and inventoried and then immediately destroyed upon a student's completion of the test.
Complete the Online Test Security Agreement	The STC must complete the Online Test Security Agreement located in the OPI MontCAS Application at https://apps.opi.mt.gov/MontCAS/frmLogin.aspx. A System Test Coordinator login is required to access the site. Copies of these agreements should be maintained in accordance with your local school district policies or the preference of the System Test Coordinator for a period of at least three school years. The OPI can request signed agreements at any time throughout the school year.

III.II Secure Handling of Printed Materials

For those students whose Individualized Education Program (IEP) or 504 Plan expresses a need for a print-on-demand of passages or items, permission for the students to receive this accommodation must first be set in TIDE at least 48 hours prior to testing. The OPI encourages ARs, STCs, and BCs responsible for inputting accommodations into student test settings to complete these at least one week prior to testing as it may take upwards of three days for changes to render in the TA interface, and in order to set any accommodation for a student in TIDE, the student's record must be flagged as having an IEP or 504 Plan. This information needs to be recorded in AIM and will be transferred to TIDE by OPI.

Printing **individual test items** for students with this accommodation will require a request to have the accommodation set for each student. For a student whose IEP or Section 504 Plan indicates a need for a paper copy of **stimuli,** access for the students to receive this accommodation must first be set in TIDE one to two weeks prior to testing by the School or District Test Coordinator.

Print requests must be approved and processed by the test administrator during test administration. The decision to allow students to use print-on-demand must be made on an individual student basis and is available only for students with an IEP or Section 504 Plan. Due to the adaptive nature of the online test, only one item can be printed at a time.

Once a student is approved to receive the printing of test items/passages, that student may send a print request to the test administrator during testing by selecting the print icon on the screen. Before the test administrator approves the student's request to print a test item/stimulus, the test administrator must ensure that the printer is on and is monitored by staff who have been trained using the Test Administrator training modules for the test. This request needs to be made for each individual item.

Destruction of printed materials and scratch paper

All test materials must remain secure at all times. When materials are not in use for testing, they must be retained securely in a locked area that can be opened only with a key or keycard by trained staff responsible for the test administration. Printed test items/passages, including embossed braille printouts and scratch paper, must be collected and inventoried at the end of each test session and then immediately destroyed according to district and/or state policies or procedures. **DO NOT** keep printed test items/passages or scratch paper for future test sessions.

Section IV. Responding to Testing Improprieties, Irregularities, and Breaches

Test security incidents, such as improprieties, irregularities, and breaches, are behaviors prohibited either because they give a student an unfair advantage or because they compromise the secure administration of the assessments. Whether intentional or by accident, failure to comply with security rules, either by staff or students, constitutes a test security incident. Improprieties, irregularities, and breaches need to be reported in accordance with the instructions in this section for each severity level. Definitions for test security incidents are provided in Table 6.

This section refers to documentation and reporting of incidents involving test security. Refer to Appendix E for a list of test security incident levels and examples of types of issues.

IV.I Impact and Definitions

Table 6: Definitions for Test Security Incidents

Type	Definition
Impropriety	An unusual circumstance that has a low impact on the individual or group of students who are testing and has a low risk of potentially affecting student performance on the test, test security, or test validity. These circumstances can be corrected and contained at the local level. An impropriety should be reported to the STC and BC immediately and entered into the OPI MontCAS Application as necessary (should an appeal be required).
Irregularity	An unusual circumstance that impacts an individual or group of students who are testing and may potentially affect student performance on the test, test security, or test validity. These circumstances can be corrected and contained at the local level. An irregularity must be reported to the STC and BC immediately and entered into the OPI MontCAS Application on the OPI website within 24 hours of the incident (should an appeal be required).
Breach	An event that poses a threat to the validity of the test. Examples may include such situations as a release of secure materials or a security/system risk. These circumstances have external implications and may result in an OPI decision to remove the test item(s) from the available secure bank. A breach incident must be reported to the BC and STC immediately and the BC should notify the OPI immediately and enter in the OPI MontCAS Application should an appeal be required.

It is important for test administrators to ensure the physical conditions in the testing room meet the criteria for a secure test environment.

Documentation of Security Incidents

All test security incidents should be recorded in the Test Security Incident Worksheet and should be maintained at the school district level. Any testing incidents that require action on the test opportunity itself, such as resetting a test, must be recorded in the TIDE Appeals system for action by the OPI State User and documented in the MontCAS Application. Refer to the *TIDE User Guide* for further guidance on how to file an appeal.

School and System (District) Test Coordinators ensure that all test security incidents are documented in a Test Security Incident Log or other tracking method. In addition to logging all test security incidents in the Test Security Incident Log, incidents requiring specific actions to be taken regarding the test itself are to be escalated to the appropriate office or agency. See Appendix E: Examples of Test Security Violations for further information.

IV.II Appeals

In the normal flow of test opportunities, students take tests in the Test Delivery System (TDS) and then submit them. Appeals are a way of interrupting this normal flow. A student may need to retake a test or have another test opportunity. A test administrator may want to invalidate a test because of a hardware malfunction or a testing incident.

Invalidating a test eliminates the test opportunity, and students have no further opportunities for the test. Resetting a test allows students to restart a test opportunity (removing all responses on the test) or allows the data entry operator to restart the data entry process. Re-opening a test allows students to return to the Montana Science Assessment, modify their answers, and resubmit the test. A grace period extension allows students to return to the questions that were locked by the 20-minute pause rule.

Test administrators should work with their STCs to submit relevant appeals in TIDE. The appeal will be reviewed by a staff member of the Montana Office of Public Instruction within 48 hours after the appeal is received.

For security incidents that result in a need to reset individual student assessments, **the request must be approved by the state**. In most instances, an appeal may be submitted to address incidents that are not security related such as reopening an assessment for a student who becomes ill and is unable to resume testing until after testing has expired.

Because the TIDE Appeals System is for action only, all appeals should also be entered in the Test Security Incident Log for school documentation purposes and, if involving a student or test, must be reported to the OPI in the required test incident online reporting system via the OPI MontCAS Application. A System Test Coordinator login is required to access the site.

Section V. Technology Infrastructure

Online testing systems used for Montana Science Assessment will be delivered via the same online platform used for Smarter Balanced assessments in English language arts and mathematics. Prior to assessment administration, Authorized Representatives (ARs), System Test Coordinators (STCs), and Building Coordinators (BCs) should review the technology infrastructure at their schools to ensure it meets the minimum requirements for administering the Montana Science assessment.

V.I Technology Resources

The Technology Guide, accessible via a link on the Montana Portal, provides Technology Coordinators with the technical specifications for online testing, including information about Internet and network requirements, hardware and software requirements, and text-to-speech functionality.

The *Operating System Support Plan* document is available for download at https://mt.portal.cambiumast.com/ under the Technology resource folder. This document provides information about supported technology and technology requirements and can be used to confirm that your school meets the minimum requirements for online testing.

There are several accessibility features embedded and supported within the test delivery system. For a complete list, refer to Appendix D. For students requiring additional assistive technology compatible with the test delivery system, including braille administrations, refer to the Assistive Technology Manual for additional configurations and support.

V.II Secure Browser

All summative assessments **must** be administered through the secure browser. The secure browser is designed to support test security by prohibiting access to external applications or navigation away from the assessment. Devices that will be used for testing must be equipped with a secure browser prior to the assessment. School Technology Coordinators are responsible for ensuring that each device to be used for testing at the school is properly secured by installing the most current secure browser.

Review *Configurations, Troubleshooting, and Advanced Secure Installation Guides* for your operating system at https://mt.portal.cambiumast.com/tech-guide.html located on the portal for more information on secure browser installation.

Section VI. General Test Administration Information

This section provides an overview of the online testing environment and guidelines for test administration. Use this section to become familiar with what students will experience in accessing the assessments, how to prepare for the assessments, and to review general rules for online testing.

VI.I Assessment Participation

Participation of Students with Disabilities and/or English Learners

Consistent with the Smarter Balanced testing plan, all students, including students with disabilities, English learners (ELs), and ELs with disabilities, should have equal opportunity to participate in the Montana Science Assessments. See the MSA UAAG for information related to universal tools, accommodations, and supports.

All students enrolled in grades 5 and 8 are required to participate in Montana science assessments. Most students in those grades will take the MSA; for those qualifying students with the most significant cognitive disabilities (approximately 1% of the student population) the Alternate Montana Science Assessment (AMSA) will be offered instead. (See https://mt.portal.cambiumast.com/alternate-science.html for information on the AMSA).

VI.II General Rules of Online Testing

This section provides a brief overview of the general test administration rules for different portions of the assessment.

Basic online testing parameters

- Students **must enter an answer for all items on a page** before going to the next page. Some pages contain multiple items. Students may need to use the vertical scroll bar to view all items on a page.
- Students may mark items for review and use the **Past/Marked Questions** drop-down list to return to those items that have already been answered within a segment. A constructed response item is considered answered if the student has taken any action within the response area. This includes random keystrokes (e.g., sdkjfdlkdjfo), one or more spaces using the spacebar, or clicking anywhere on a Grid Item – Hot Spot Item, etc.

Pause Rules

Students are logged out of the test when they click the Pause button and then click Yes to pause their test. Students will need to have access to their login information (name, SSID, and test session) after pausing their test. The student will be required to log back into the student interface when they are ready to continue testing. The Montana Science Assessment may be administered in multiple sessions across more than one day.

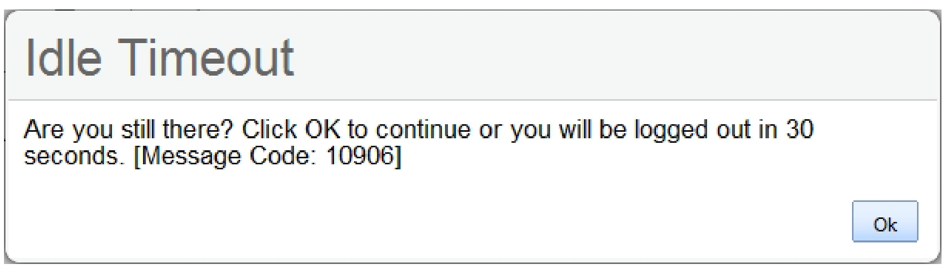
- Any highlighted text will be saved regardless of how long the test is paused so long as the student is logging back into a device of the same operating system. Notes on the digital notepad will not be saved when a test is paused, regardless of how long the test is paused.

- In the event of a technical issue (e.g., power outage or network failure), students will be logged out and the test will automatically be paused. The students will need to log in again upon resuming the test.
- There will be a 20-minute pause rule set for the Montana Science Assessment. If a student pauses the test for more than 20 minutes, the student will not be able to return to the questions that they previously answered. If a student pauses on a page that contains multiple questions, the student will be able to return to the same page and respond to the remaining questions even after the 20-minute pause.
- If the Montana Science Assessment is administered in one day, a break of 10 minutes is recommended between two sessions, with each session being approximately 45 minutes long.

Test timeout (due to inactivity)

As a security measure, students and test administrators are automatically logged out of the test after 30 minutes of inactivity. *Activity* is defined as selecting an answer or navigation option in the assessment (e.g., clicking [**Next**] or [**Back**] or using the **Past/Marked Questions** drop-down list to navigate to another item). Moving the mouse or clicking on an empty space on the screen is not considered activity.

Before the system logs the student out of the assessment, a warning message will be displayed on the screen. If students do not click [**Ok**] within 30 seconds after this message appears, they will be logged out. Clicking [**Ok**] will restart the 30-minute inactivity timer.



VI.III Testing time and scheduling

Testing Window

Each district will establish a schedule for the administration of the Montana Science Assessment using a testing window as defined below:

Grades 5 and 8:

- The online test window is March 6 – May 26, 2023.
- Testing may continue up to and including the last day of the testing window.

A system or school may establish more specific windows within the constraints of the defined window described above.

Scheduling Time for Testing

The Montana Science Assessment is untimed. The recommended testing time is at least 90 minutes plus an additional 10 minutes for reading the directions to the students. The recommended testing time is an estimate of how much time students will need to complete the Montana Science Assessment. Students should be given additional time if it is needed. Ninety minutes is the minimum time that must be provided to students unless all students in a testing room have finished early.

VI.IV Establishing Appropriate Testing Conditions

Building Coordinators (BCs), Test Administrators (TAs), and Proctors (PR) will need to work together to determine the most appropriate testing option(s) and testing environment based on the number of computers available, the number of students in each tested grade, and the estimated time needed to complete each test. Testing students in classroom-sized groups is preferable. Establishing classroom-sized groups reduces test fear and anxiety for the students and facilitates monitoring and control for the test administrator.

Establish procedures to maintain a quiet testing environment throughout the test session, recognizing that some students will finish more quickly than others. If students are allowed to leave the testing room when they finish, explain the procedures for leaving without disrupting others and where they are expected to report once they leave. If students are expected to remain in the testing room until the end of the session, instruct them on what activities they may engage in after they finish the test.

Section VII. Universal Tools, Designated Supports, and Accommodations

Appendix D includes a list of the universal tools, designated supports, and accommodations available to students for the Montana Science Assessment. Student eligibility for accommodations and supports is communicated by district-level users through Achievement In Montana (AIM). OPI uploads the student data file from AIM into the Test Information Distribution Engine (TIDE). TIDE then distributes this information to the appropriate system.

The *Montana Usability, Accessibility, and Accommodations Guidelines* are intended for school-level personnel and decision-making teams, including Individualized Education Program (IEP) and Section 504 teams, as they prepare for and implement the Montana Science Assessment. The *Guidelines* provide information for classroom teachers, English language development educators, special education teachers, and instructional assistants to use in selecting and administering universal tools, designated supports, and accommodations for those students who need them. The *Guidelines* are also intended for assessment staff and administrators who oversee the decisions that are made in instruction and assessment.

If school or district staff members identify a designated support and/or accommodation that they believe should be offered but is not available, the school or district should provide that information to the District Test Coordinator, who will inform OPI of the request.

Table 7: Definitions for Universal Tools, Designated Supports, and Accommodations

Type	Definition
Universal Tools	Access features of the assessments that are either provided as digitally delivered (embedded) components of the test administration system or are separate (non-embedded) from it. Universal tools are available to all students based on student preference and selection.
Designated Supports	Access features of the assessments available for use by any student for whom the need has been indicated by an educator (or team of educators working with the parent/guardian and student). They are either provided as embedded components of the test administration system or are non-embedded.
Accommodations	Accommodations are changes in procedures or materials that increase equitable access during the Montana Science assessments. Assessment accommodations generate valid assessment results for students who need them; they allow these students to show what they know and can do. Accommodations are available for students with documented IEPs or 504 Plans. State-approved accommodations do not compromise the learning expectations, construct, grade-level standard, or intended outcome of the assessments. They are either provided as embedded components of the test administration system or are non-embedded.

State users, ARs, STCs, and BCs have the ability to set embedded and non-embedded designated supports and accommodations. **Designated supports and accommodations must be set in TIDE prior to starting a test session.**

For additional information about the availability of designated supports and accommodations, refer to the *Usability, Accessibility, and Accommodations Guidelines* for complete information. For information on updating student settings, refer to the *TIDE User Guide*. The *Test Administrator User Guide* also contains information on how to use some of these settings. These user guides can be found on the portal https://mt.portal.cambiumast.com/.

Section VIII. Prior to Test Administration

VIII.I Tasks to Complete Prior to Administration

- **Authorized Representatives (ARs), System Test Coordinators (STCs), Building Coordinators (BCs) Test Administrators (Tas) and Proctors (PR) should verify that students are provided the opportunity to complete the Practice Test prior to testing.** This allows students the opportunity to become familiar with all item types, the universal tools, buttons, and any allowable designated supports and/or accommodations.
- **ARs and BCs should ensure all TAs and PRs have login information for the TA interface.**
- **ARs, BCs, TAs and PRs should verify and/or update student demographic information and test settings.**

  Please make sure that students are correctly assigned to their district, school, and grade in AIM. The state will upload students into TIDE.

  In addition to the correct school and grade, BCs, TAs, and PRs should verify that all students have accurate test settings in TIDE, including designated supports and accommodations (embedded and non-embedded) for each content area prior to testing.

  Please ensure that any changes are placed in AIM and those will be reflected in TIDE within 24 to 48 hours. ARs, STCs, BCs, and TAs can set embedded and non-embedded designated supports and accommodations in TIDE for students who require them. PRs may view student information; however, PRs cannot add, upload, or modify student information. Within the TA Interface, a PR can turn off universal tools prior to the start of the test.

  IMPORTANT: Any additions/modifications/deletions of students and changes to student test settings must be completed **before the student can test**. The update, once made, may take **up to 24 hours to appear in the TA interface**. Failure to correct test settings before testing could result in the student's not being provided with the needed accommodations and/or designated supports at the time of testing. This is considered a testing irregularity.

  For information on how to view and modify student information, please refer to the *TIDE User Guide*.

  It is important for anyone with access to student information to remember that student personal information, including the student's SSID, is confidential. **If materials containing student personal information and/or SSIDs are distributed to students, these materials must be collected before the students leave the testing room and either securely stored to be used in a subsequent test session or shredded.** For additional information about security protocols, refer to section *Appendix E*.

- **BCs should ensure that test administrators have necessary student login information.**

  Students will log in to the online testing system using their preferred first name, SSID, and a test session ID. Prior to starting a test session, test administrators must have a record of each student's first name and SSID. This information must be provided to each student to complete the login process. It is suggested that the SSID and the student's first name (as it appears in TIDE) be printed on a card or piece of paper and distributed to students just prior to testing to help them type it in the computer accurately. Student information is confidential; therefore, the cards/papers with this information must be kept secure until used during a test session. Test administrators may provide additional identifying information on the printed card. The loss of login information is considered a security breach at the local level and needs to be reported for test security reasons.

When a test administrator creates a test session, a unique session ID is randomly generated. This session ID must be provided to students before they log in. Please refer to the *Test Administrator User Guide* for more detailed information on how to obtain session IDs.

- **BCs should ensure that test administrators have any additional required resources.**

Section IX. Day of Test Administration

Use the following information and script to assist students with the login procedures. Please refer to the *Test Administrator User Guide* at https://mt.portal.cambiumast.com/ to become familiar with the Online Testing System. The Test Administrator (TA) or Proctor (PR) should verify the security of the testing environment prior to beginning a test session. Test administrators must ensure that students do not have access to non-approved digital, electronic, or manual devices during testing.

To ensure that all students are tested under the same conditions, the test administrator should adhere strictly to the script for administering the test unless a student is assigned the non-embedded designated support or accommodation for Simplified Test Directions. The script can be found in the blue shaded boxes on the following pages, indicated by the word "SAY". They should be read exactly as they are written using a natural tone and manner. If the test administrator makes a mistake in reading a direction, the test administrator should stop and say, "I made a mistake. Listen again." Then the direction should be reread. When asked, the test administrator should answer questions raised by students but should never help the class or individual students with specific test items. **No test items can be read to any student for any content area, unless the student is assigned a Read Aloud non-embedded designated support or accommodation as described in the *Montana Usability, Accessibility, and Accommodations Guidelines*.**

If the class is resuming a test and the test administrator is sure that all students are able to log in without hearing the login directions again, the test administrator may skip to the directions indicated in section Testing Over Multiple Sessions or Days.

The test administrator should try to maintain a natural classroom atmosphere during the test administration. Before each test begins, the test administrator should encourage students to do their best. Test administrators are responsible for monitoring student progress for the Montana Science Assessment.

IX.I Administering a Test Session

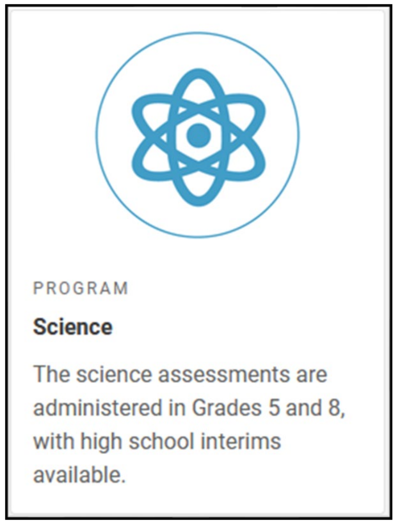
The test administrator should follow the steps outlined in this section to administer a test session.

The test administrator must create a test session before students can log in to the Student Testing System **(but no more than 30 minutes prior or the system will time out).** When a test administrator creates a test session, a unique session ID is randomly generated. This session ID must be provided to the students before they log in and should be written down. Test administrators should follow these steps to create a session:

1. The test administrator logs in to the TA Interface.

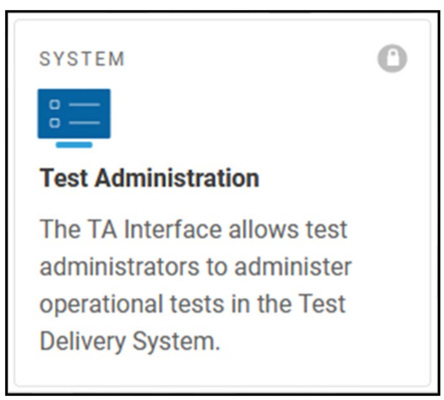
a) From the Montana Comprehensive Assessment Program Portal (https://mt.portal.cambiumast.com/) select the Science Assessment program card.

Figure 1: Science Assessment Card



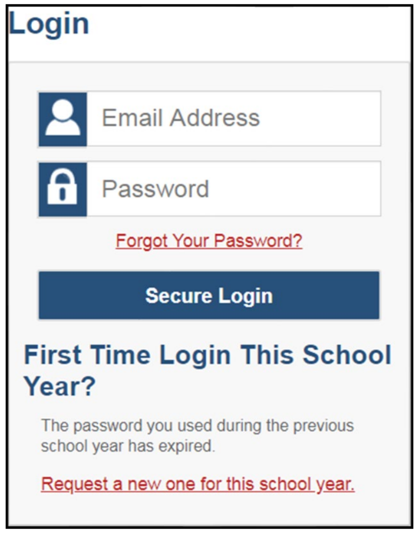
b) Select the **Test Administration** card.

Figure 2: Test Administration Card



c) On the *Login* page, the Test Administrator enters their email address and password associated with their TIDE account in the respective fields. Then, the Test Administrator clicks **Secure Login**. The TA Interface appears.

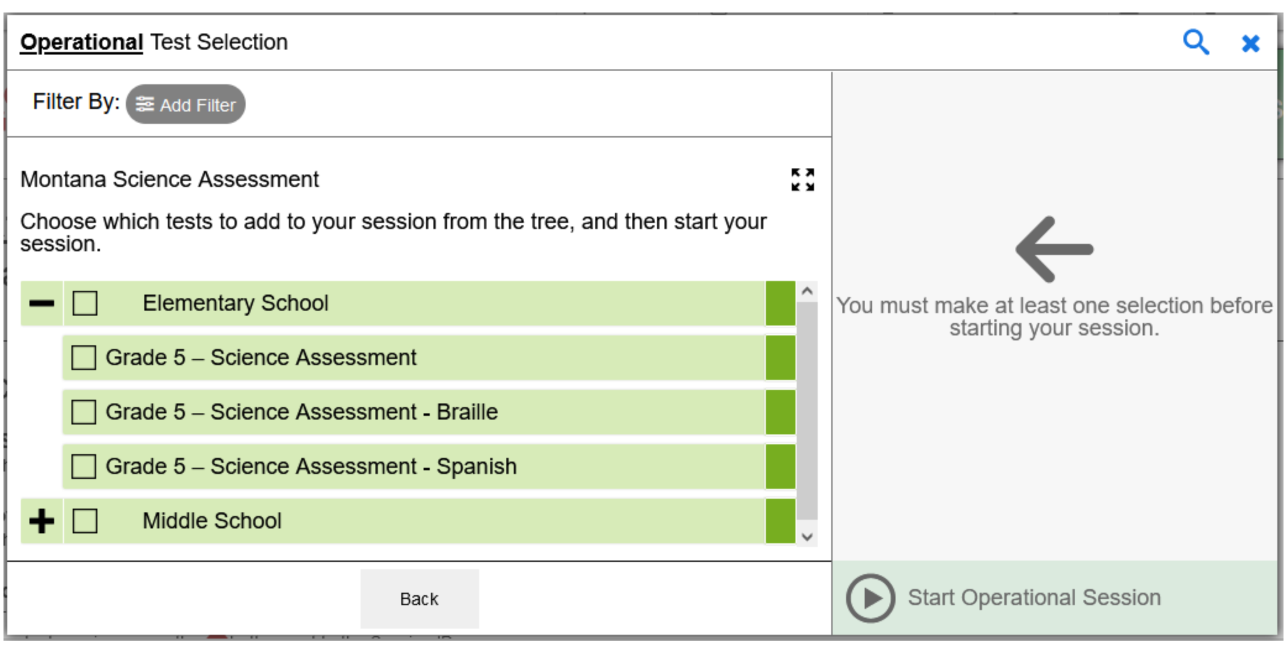
Figure 3: Login Page



2. The test administrator begins the test session.

When a Test Administrator logs in to the TA Interface, the *Test Selection* window opens automatically. Afer clicking the assessment(s) to be administered during the test session, the test administrator clicks on the [Start Operational Session] button to begin the test session and generate the session ID that students use to join that test session. The system-generated session ID appears in the Operational Session ID box at the top of the TA interface screen.

Figure 4: Test Selection Window



3. The test administrator prepares students for testing and login.

Figure 5: Operational Session ID

SAY: Today, you will take the [NAME OF TEST (i.e., Grade 5 - Science)] test. You will be given a test session ID that is required to start the test. Before logging in, let's go over some test rules.

You must answer each question on the screen before clicking NEXT. If you are unsure about an answer, provide what you think is the best answer. Some questions may require more than one answer or may require different kinds of interactions. If you would like to review that answer at a later time, mark the question for review before going on to the next question. You may go back and change the answer during this test session.

You may pause at any point in the test by clicking PAUSE rather than NEXT after answering a question. Please raise your hand if you need a break and ask permission before clicking PAUSE.

SAY: If you pause your test for more than 20 minutes, you will not be able to go back and change your answers to any question, even the ones you marked for review. You may check your answers to any questions before you pause your test.

SAY: Your answers need to be your own work. Please keep your eyes on your own test and remember, there should be no talking. If you have a cell phone, smartwatch, or other non-approved electronic devices, please raise your hand and I will collect it before the test begins. I will return your device at the end of this test session.

If you finish your test early, please raise your hand and sit quietly.

Give students the test session ID and other log-in information. The test session ID, the statewide student identifier (SSID), and the student's first name may be provided to students ahead of time on a card or piece of paper to help them type it in the computer accurately. Student information is confidential; therefore, the cards/papers with this information must be collected after each test session and securely destroyed after testing is complete. **The test administrator should write down the session ID for his or her own records, in case they get involuntarily logged out of the system.** The test administrator will be logged out of the session if there is no activity for 30 minutes by the TA or a student. Having the session ID will allow the test administrator to resume the session.

Test administrators may also write the test session ID on a classroom dry-erase or chalk board or another place where students can see it. Make sure students know that the test session ID must be entered **exactly** as it is written without extra spaces or characters.

When providing students with the test session ID number, read aloud the directions below.

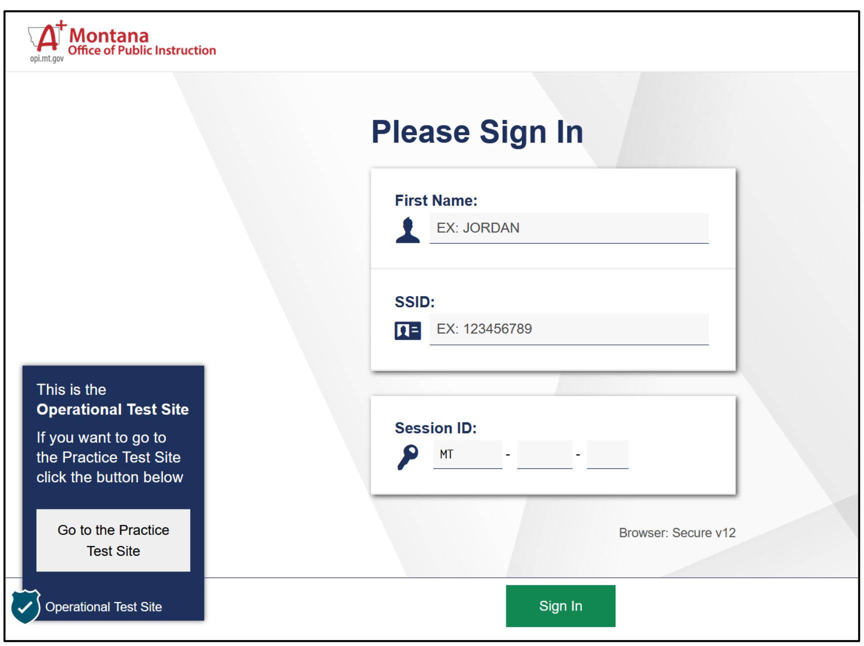
SAY: Now we are ready to log in. Once you have logged in, you will have to wait for me to approve the test before you start. I'll be checking that you have correctly entered the test session ID and other information.

Enter your legal first name, not your nickname, followed by your SSID number. Then enter the test session ID. Raise your hand if you need help typing this information on your keyboard.

Now click SIGN IN. Once you have successfully logged in, you will see a screen with your first name and other information about you. If all of the information on your screen is correct, select YES to continue. If any of the information is incorrect, please raise your hand and show me what is incorrect.

Ensure that all students have successfully entered their information. The test administrator should ensure that the students use their legal first names as they appear in TIDE, not nicknames. If a student is unable to log in, they will be prompted to try again and provided with a message describing the reason (an invalid SSID, for example). If the student is still having difficulty, the test administrator can look up the correct information using the Student Lookup function on the TA Interface or in TIDE. Test administrators may assist students with logging in if necessary.

Figure 6: Student Login Page



4. *The test administrator confirms student test selection.*

> **SAY:** On the next screen, select the [INSERT NAME OF TEST (i.e., Grade 5 Science)] test, and then click START TEST. After you have selected your test, you will see a screen with a moving bar and message saying that you are waiting for test administrator approval. Please wait quietly while I check and approve each of your tests.

5. *The test administrator views and approves students who are waiting for test session approval.*

After students have selected a test, the test administrator verifies that each student selected the appropriate test before approving that student for testing. **It is very important that the test administrator pays close attention to the test name prior to approving to be sure it is the correct test to be administered at that time.** To do this:

a. Select the [**Approvals (#)**] button.

Figure 7: Students Waiting for Test Session Approval in the TA Interface

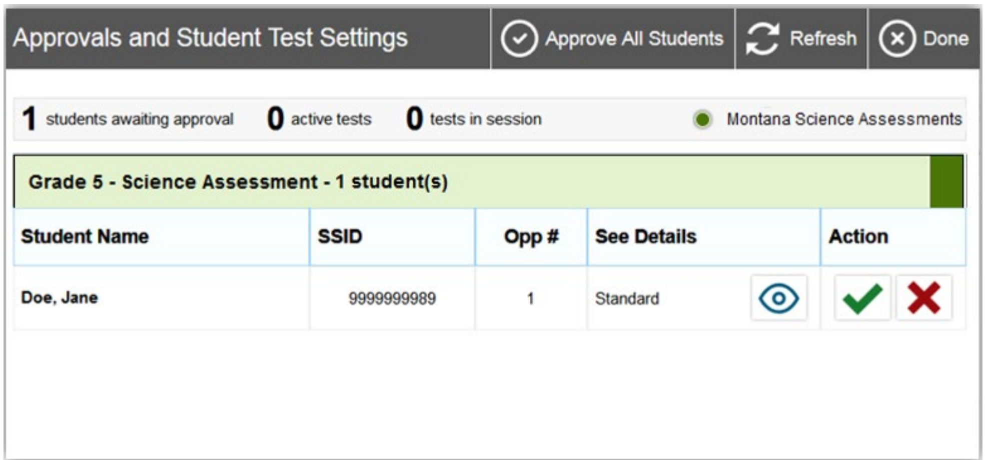
b. A new window opens that shows a list of students, organized by test name. The test administrator should review the list to ensure that students are taking the correct test (grade level), and the correct content area (Science).

c. The test administrator should also review the test settings assigned to each student to ensure that they are correct.

   i. If a student's settings are incorrect, do *not* approve that student to begin testing. The test administrator will need to work directly with an AR or STC to correct the test settings in TIDE before approving the student to begin testing. *(Reminder: It may take up to 24 hours for changes to appear in the TA interface.)* **Ensuring the test settings are correct before the student begins testing is critical to avoid the need for a test reset that may result in additional testing for the student later.**

d. When the correct test is selected and test settings are verified, the test administrator clicks [**Approve**] or [**Approve All Students**].

Figure 8: Approvals and Student Test Settings Window

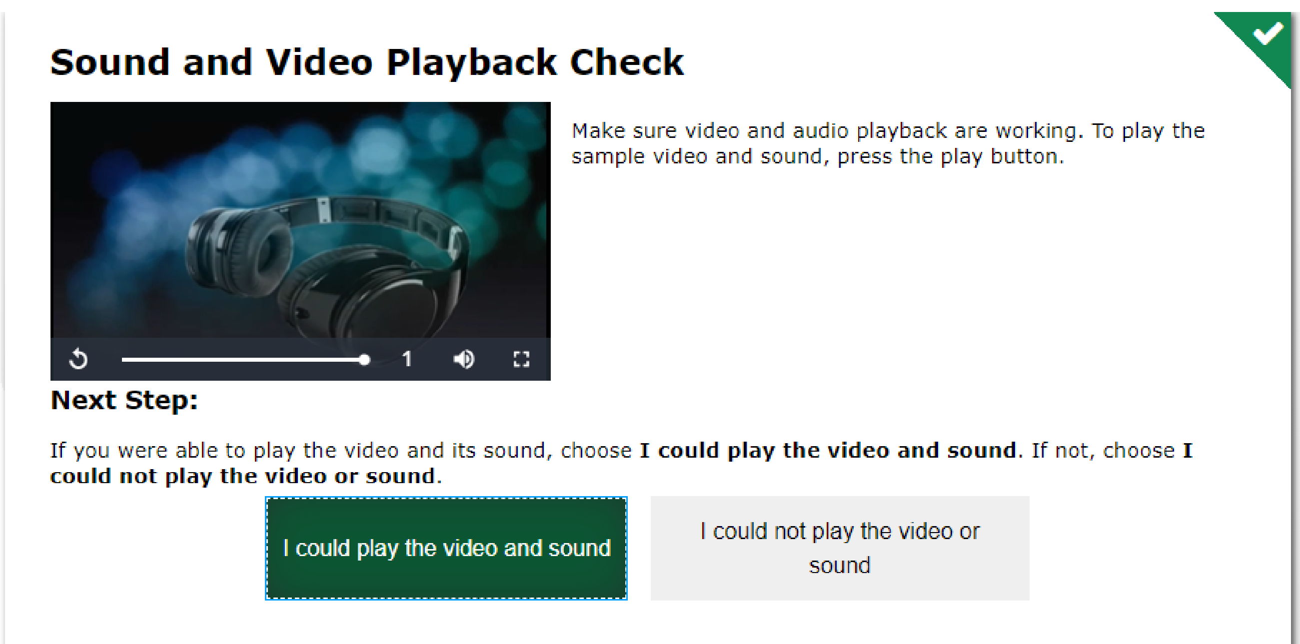


SAY:	You will now see a screen asking you to check your test content area and settings. If all the information is correct, you may select YES, START MY TEST. If any of it is incorrect, please raise your hand.

Each student may be logging in at a different time. The test administrator should monitor the sessions and log in all students who are currently ready before assisting any students who are having problems.

The students will also see a video playback check, as some of the Montana Science Assessment items include short video simulations. There are no audio items on the Montana Science Assessment (unless a student requires text-to-speech designated support).

Figure 9: Sound and Video Playback Check Page



For ALL students read the following:

SAY: The questions in this test cover many science topics. Although some of the questions may look difficult or unfamiliar, you may be able to figure out the answers if you read and think about them carefully.

If you are unsure of an answer, provide what you think is the best answer; there is no penalty for guessing. If you would like to review that answer at a later time, mark the question for review. Mark for Review is available in the upper right corner of the answer and represented by three horizontal lines. Flagging the question will remind you to go back and decide whether you want to change the answer later.

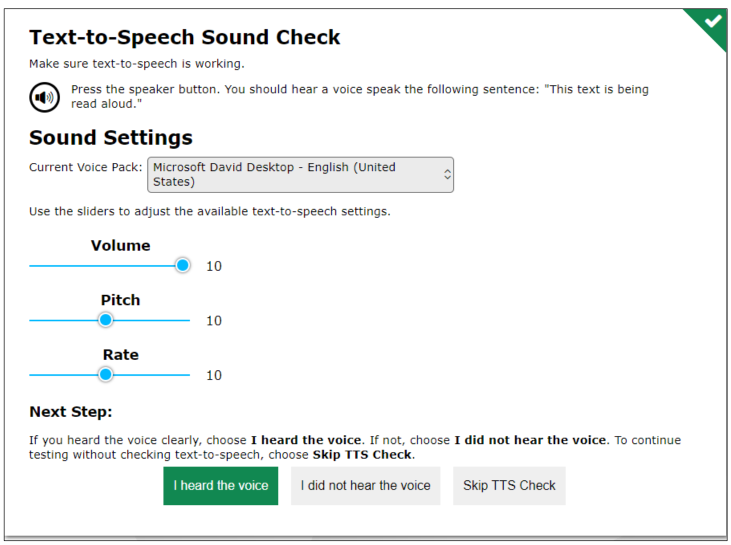
SAY: Your answers need to be your own work. Please keep your eyes on your own test and remember that there should be no talking.

When you are ready to begin your test, click [BEGIN TEST NOW] at the bottom of the page.

Text-to-Speech Designated Support

Students marked as requiring text-to-speech designated support in TIDE will be asked to verify that their headsets are working properly.

Figure 10: Text-to-Speech Sound Check Page



If a student's volume is not working, troubleshoot basic steps such as ensuring that the headset is fully plugged in, mute is not inadvertently selected, and the volume setting is turned up. If the student is still experiencing issues, contact your School Technology Coordinator or the Montana Help Desk at 1-877-365-7915 or mthelpdesk@cambiumassessment.com.

For only those students requiring text-to-speech designated support, read the following.

SAY: You will also see a Text-to-Speech Sound Check below the Audio/Video checks. Check the prompts to verify the sound on your computer is working. Put your

headsets on and click the green icon of the speaker to hear the sound. If you hear the voice, click [I heard the voice]. If not, raise your hand. You can adjust the volume, pitch, and rate of the voice by using the sliders.

6. The test administrator monitors student progress.

Monitoring Test Progress

Once students have started their tests, the test administrator should circulate through the room to ensure that all conditions of test security are maintained. If the test administrator witnesses or suspects the possibility of a test security incident, the AR and STC should be contacted immediately in accordance with the security guidance provided in this manual.

The test administrator may also use the TA Interface to view the testing progress of any student. This site will not show test items or scores but will let the test administrator see overall progress of each student (e.g., 40%).

Figure 11: TA Refresh Button

While the TA Interface is designed to automatically refresh every minute, the test administrator can refresh it manually at any time by clicking the button at the top left of the page. Do NOT click the web browser's refresh button to refresh the TA Interface.

If the test is administered on the same day, with a 10-minute break between 45-minute sessions, use the script below.

When there are 10 minutes left in the first test session, read the following.

SAY: You have 10 minutes left. Please review any completed or marked-for-review questions now. Please PAUSE the test when you are finished reviewing your work. You will be able to return to the test after the break.

If the test is administered over two days, use the script below.

When there are 10 minutes left in the first test session, read the following.

SAY: You have 10 minutes left. Please review any completed or marked-for-review questions now. Please PAUSE the test when you are finished reviewing your work. You will be able to return to the test when the second session of the test

is administered. You will not be able to go back to the items you responded to today.

If the pause is longer than 20 minutes, and the original test session has expired, the test administrator will need to start a new test session. In addition, if the test is paused for more than 20 minutes, students will not be allowed to review previously responded items.

Student Directions During Testing

If you notice that a student is off task, you may read the statement below **verbatim**.

SAY: It is important that you do your best. Do you need to pause the test and take a break?

If a student is concerned about an item, you may direct the student to review a tutorial by reading the script below **verbatim**.

SAY: Try your best and choose the answer that makes the most sense to you. If you are unsure about how a question works, you can review a tutorial by clicking on the “i" button from the question context menu.

IMPORTANT:

- If the test administrator is using the TA Interface and navigates to TIDE or the TA Practice or Training Site, the session may stop, and all students in the session will be logged out. The session cannot be resumed. A new session will have to be created, and the students will have to log in to the new session to resume testing. When starting a new session, give the new session ID to the students so they can log in and resume testing.
- If the TA Interface or TA training site browser is accidentally closed while students are still testing, the session will remain open until it times out after 30 minutes. The test administrator can open the browser and navigate back to the TA Interface. The test administrator will be prompted to re-enter the active session ID.
- As a security measure, test administrators are automatically logged out of the TA Interface after 30 minutes of test administrator user inactivity and student inactivity in the test session, which will result in closing the test session. If this occurs, the test administrator will have to create a new session and the students will have to log in to the new session to resume testing. When starting a new session, the test administrator should give the students the new session ID so that they can log in and resume testing.

7. The test administrator ends the test session and logs out of the Test Administrator Site.

The information button, , in the TA Interface also allows test administrators to monitor and understand reasons that halt student progress.

When there are approximately ten minutes left in the test session, the test administrator should give students a brief warning.

SAY: We are nearing the end of this test session. Please review your answers, including any questions you marked for review now. Do not submit your test unless you have answered all of the questions.

[The following should be read for all tests]

SAY: If you need additional time let me know.

After answering the last item in the test, each student is presented with a screen prompting the student to review answers (marked and unmarked) for all items available to the student prior to submitting the test. A test cannot be paused after the last item has been presented to the student.

After answering the last question, students must submit their tests. If students would like to review their answers before submitting their test, they should click [**REVIEW MY ANSWERS**] and then [**SUBMIT TEST**] after they finish reviewing. Once a student clicks [**SUBMIT TEST**], the student will not be able to review answers.

SAY: This test session is now over. If you have not finished, click PAUSE, and you will be able to finish at another time.

Follow appropriate testing policies regarding allowing students who need additional time to finish testing. Keep in mind the 20-minute pause rule; students whose test is paused for longer than 20 minutes will not be able to go back to any questions that contain a response or are marked for review. Students will also not have access to notes entered in the Digital Notepad.

SAY: Remember, if you are not finished and need to pause your test for more than 20 minutes, you will not be able to go back to the questions from this test session, including questions marked for review. You will also not have access to the notes you entered in the Digital Notepad.

SAY: If you have answered all the questions on your test and have finished reviewing your answers, click SUBMIT TEST. I will now collect any scratch paper or other materials.

Test administrators should click [**Stop Session**] to end the test session and pause any student test in the session that is still in progress. When finished, test administrators can log out of the TA Interface by clicking the [**Logout**] button at the top right. Test administrators should also collect any scratch paper or material with student information used to assist with logging into the test.

Testing Over Multiple Sessions or Days

If the test administrator intends to administer the test over the course of multiple days for a student or group of students, test administrators may ask students to pause after they reach a designated point. There is nothing built into the system to prevent students from progressing from one section of the test to another. The test administrator should give the students clear directions on when to pause. For example, test administrators may designate a certain amount of time for testing. It is advised that students complete all parts of the item on the page, such as an item cluster, before pausing the test. This guidance may be written on a dry-erase board, chalkboard, or another place that students can easily see.

When testing is resumed on a subsequent day, the test administrator will need to start a new test session and provide a new session ID.

Use the script below the following day when the next session is administered.

SAY:	**Enter your first name followed by your student ID number using the information on the paper I gave you. Then enter the test session ID number exactly as it is written without extra spaces or characters. Raise your hand if you need help typing this information. Follow the on-screen prompts and sign into the test session. Once I approve you for the test, you can resume working on the test.**

Section X. After Science Assessment Administration

Destroying Test Materials

Federal law—the Family Educational Rights and Privacy Act—prohibits the release of any student's personally identifiable information. Any printed materials must be securely stored and then shredded.

As a reminder, any printed and paper test materials must be securely shredded immediately following each test session and may not be retained from one test session to the next.

Reporting Testing Improprieties, Irregularities, and Breaches

Throughout testing, ensure that all test security incidents were reported in accordance with the guidelines in Ensuring Test Security and Responding to Testing Improprieties, Irregularities, and Breaches in this manual.

Appendix A. Frequently Used Terms

Table 8: Frequently Used Terms

Term	Definition
Accommodations	Changes in procedures or materials that increase equitable access to the Montana Science Assessments. Assessment accommodations generate valid assessment results for students who need them; they allow these students to show what they know and can do. Accommodations are available for students with documented Individualized Education Programs (IEPs) or 504 Plans. OPI-approved accommodations do not compromise the construct, grade-level standards, or intended outcome of the assessments. See the *Montana* *Usability, Accessibility, and Accommodations Guidelines* for complete information.
Appeal	Authorized users may submit and view requests for reopening students' assessments in accordance with state policy. These requests must result from a test security incident or incorrect test setting that impacted testing. All requests must be approved by a state education agency representative.
Authorized Representative (AR)	District-level staff member who has the same capabilities as a State user in TIDE. ARs may be assigned by the State. If assigned, an AR can upload, add, modify, and remove student records. All other responsibilities for this role should be assigned by the State.
Break	There is no limit on the number of breaks or the length of a break that a student may be given according to his or her unique needs. However, breaks of more than 20 minutes will prevent the student from returning to items already attempted, completed, and/or flagged for review by the student.
Building Coordinator (BC)	School staff member responsible for monitoring the test schedule, process, and TAs. BCs are also responsible for ensuring that TAs have been appropriately trained and that testing is conducted in accordance with the test security and other policies and procedures established by the Montana Office of Public Instruction.
Designated Supports	Accessibility features of the assessments available for use by any student for whom the need has been indicated by an educator (or team of educators working with the parent/guardian and student). See the *Montana* *Usability, Accessibility, and Accommodations Guidelines*, for complete information.
Item	A test question or stimulus presented to a student to elicit a response.
Item Cluster	A test question that presents a phenomenon with multiple parts associated.
Pause	Action taken by a student or test administrator to temporarily halt the test during any part of the test, as needed. Pauses of more than 20 minutes in the test will prevent the student from returning to items already attempted, completed, and/or flagged for review. More information on test pausing is available in section General Rules of Online Testing.
Proctor (PR)	District or school personnel responsible for administering the MontCAS assessments in a secure manner in compliance with the policies and procedures outlined in the *Test Administration Manual*.
Reopen	A specific appeal in the TIDE Appeals module. It applies to a test that has already been submitted or has expired and allows the student to access the previously closed test. For example, a test may be reopened if a student started a test and became ill and was unable to resume testing until after the testing opportunity expired. Permission for a reopen is initiated through the Appeals process in TIDE. See the *TIDE User Guide* for more information on the Appeals process.

Term	Definition
System Test Coordinator (STC)	District-level staff member who is responsible for the overall administration of the summative assessment in a district. STCs should ensure that the Building Coordinators (BC) and test administrators in their districts are appropriately trained and aware of policies and procedures. In the event there is no STC, another designated individual will be assigned these responsibilities by the State.
Test Administrator (TA)	School staff member responsible for monitoring the test schedule, process, and PRs. TAs are also responsible for ensuring that PRs have been appropriately trained and that testing is conducted in accordance with the test security and other policies and procedures established by the Montana Office of Public Instruction.
Secure Browser	A web browser that is downloaded and installed on a computer prior to a student beginning the test. Each browser is specifically designed for use with one or more of the MontCAS assessments to provide secure access to the test for each content area and prevent students from accessing specific hardware and software functions (e.g., other browsers, screenshot programs) that are not allowed during the assessments.
Session	A timeframe in which students actively test in a single sitting. The length of a test session is determined by building or district administrators who are knowledgeable about the periods in the building and the timing needs associated with the assessment.
Statewide Student Identifier (SSID)	A statewide, unique student identifier assigned to each student by his or her state education agency. This is the same identifier used on other state assessments.
Stimulus/Stimuli	Material or materials used in the test context which form the basis for students to answer connected test items. Many items/tasks for the assessments include a stimulus along with a set of questions to which the student responds. Examples of stimuli include: texts viewed on a computer screen, images with audio presentations for students to listen to, simulated web pages for students to use for research or scenarios, charts and graphs.
Test Security Incident Log	Document of record for a district to record all test security incidents.
Testing Breach	A security event that poses a threat to the validity of the test. Examples may include such situations as a release of secure materials or a security/system risk. These circumstances have external implications and may result in a decision to remove the test item(s) from the available secure bank. A breach incident must be reported immediately to the state level. For specific details on how to proceed when an incident has occurred, please refer to Section IV of this manual. Also see the OPI Test Security webpage.
Testing Impropriety	An unusual circumstance that has a low impact on the individual or group of students who are testing and has a low risk of potentially affecting student performance on the test, test security, or test validity. These circumstances can be corrected and contained at the local level. For specific details on how to proceed when an incident has occurred, please refer to Section IV of this manual. Also see the OPI Test Security webpage.
Testing Irregularity	An unusual circumstance that impacts an individual or group of students who are testing and may potentially affect student performance on the test, test security, or test validity. These circumstances can be corrected and contained at the local level but are submitted in the online system for resolution of the Appeal for testing impact. For specific details on how to proceed when an incident has occurred, please refer Section IV of this manual. Also see the OPI Test Security webpage.
Universal Tools	Available to **all** students based on student preference and selection. See the *Usability, Accessibility, and Accommodations Guidelines* for complete information.

Appendix B. Secure Browser for Testing

Test administrators should be aware of the following guidelines and employ the necessary precautions while administering online assessments:

- **Close External User Applications**
  Prior to administering the online assessments, test administrators should check all computers that will be used and close all applications except those identified as necessary by the Building Coordinator. After closing these applications, the test administrator should open the secure browser on each computer.

  The secure browser and Student Interface automatically detect certain applications that are prohibited from running on a computer while the secure browser is open. The secure browser will not allow a student to log in if the computer detects that a forbidden application is running. A message will also display that lists the forbidden application(s) that needs to be closed.

If a forbidden application is launched in the background while the student is already in a test, the student will be logged out and a message displayed. The student will have to close the forbidden application, reopen the secure browser, and log in to continue working on the assessment. This would be considered a test security incident.

- **Testing on Computers with Dual Monitors**
  Systems that use a dual monitor setup typically display an application on one monitor screen while another application is accessible on the other screen.

  In most cases, students should *not* take online assessments on computers that are connected to more than one monitor. The only exception should be in extremely rare circumstances such as when a TA is administering a test via read-aloud and wants to have a duplicate screen to view exactly what the student is viewing for ease of reading aloud. In these rare cases where a dual monitor is allowed, MSA requires that the monitors be set up to "mirror" each other. Building Coordinators can assist test administrators in setting up the two monitors to ensure they mirror each other rather than operate as independent monitors.

  In these cases, all security procedures must be followed, and the test administered in a secure environment to prevent others from hearing the questions or viewing the student or test administrator screens.

Secure browser error messages

Possible error messages displayed by the open-source secure browser are shown in Table 9.

Table 9: Secure Browser Error Messages

Message	Description
Secure Browser Not Detected	The system automatically detects whether a computer is using the secure browser to access the online testing system. The system will not permit access to a test other than through the secure browser; under no condition may a student access a test using a non-secure browser.
Unable to Establish a Connection with the Test Delivery System	If a computer fails to establish a connection with the test delivery system, a message will be displayed. This is most likely to occur if there is a network-related problem. The causes may include an unplugged network cable or a firewall not allowing access to the site.

Force-quit commands for secure browsers

In the rare event that the Secure Browser or test becomes unresponsive, you can force-quit the Secure Browser. Please note that the Secure Browser hides features such as the Windows task bar or Mac OS X dock. If the Secure Browser is not closed correctly, then the task bar or dock may not reappear correctly, requiring you to reboot the device. Avoid using a force-quit command if possible.

To force the Secure Browser to close, use the keyboard command for your operating system as shown below. This action logs the student out of the test. When the Secure Browser is opened again, the student logs back in to resume testing.

Table 10: Force Quit Secure Browser Keyboard Commands

Operating System	Key Combination
Windows*	• **Ctrl** + **Alt** + **Shift** + **F10**
Mac OS X*	▪ **Ctrl** + **Alt** + **Shift** + **F10**. The **Ctrl** key may appear as **Control**, **Ctrl**, or **^**
Linux	• **Ctrl** + **Alt** + **Shift** + **Esc**

* If you are using an Apple keyboard, you may need to press **Ctrl** + **Shift** + **Option** + **F10.** If you are using a laptop or notebook, you may also need to press **Function** before pressing **F10**.

Force-quit commands do not exist for the Secure Browser for iOS, and Chrome OS devices.

You are strongly advised against using the force-quit commands, as the secure browser treats this action as an abnormal termination. The secure browser hides features such as the Windows taskbar. If the secure browser is not closed correctly, then the taskbar may not reappear correctly. These commands should be used only if the [**Close Secure Browser**] button does not work.

Appendix C. Practice Test

In preparation for the test and to expose students to the various item-response types in the Montana Science Assessment, it is highly recommended that all students access the Practice Tests, available at https://mt.portal.cambiumast.com/science.html. Each resource offers students a unique opportunity to experience a test in a manner like what they will see on the MSA and prepares students for testing. These resources will not be automatically scored; however, the ability to score each item is embedded within the online practice test.

Overview of the Practice Tests

The Montana Science Practice Tests allow educators and students to experience a full grade-level test and gain insight into the assessments. The Practice Tests mirror the assessments, but responses are not saved and are not scored. Both the Grade 5 and Grade 8 test include a variety of item response types and difficulty levels—the Practice Test also includes a comprehensive reflection of embedded universal tools, designated supports, and accommodations, which should also be provided to students with any non-embedded universal tools, designated supports, and accommodations as allowed on the summative assessments.

Test Administrator login for the Practice Test

The Test Administrator Training Test site provides an opportunity for Test Administrators and Proctors to practice setting up a test session using a Practice Test. Access to the Test Administrator Training Test site requires specific login credentials provided by the State. This site may only be used by authorized users with a TIDE account.

For additional information about how to set up a Practice Test session, refer to the Test Administrator User Guide located at https://mt.portal.cambiumast.com/.

Do NOT use the live Student Interface or TA Interface for practice. Doing so constitutes a test security incident (breach). For all Training Test sessions, use the TA Training site and the Practice Tests or Training Tests.

It is highly recommended that all students access the Practice Test site before taking a test. Teachers are encouraged to conduct a group walk-through of the Practice Test with their students to promote familiarity with the testing format and basic test rules and to allow for free and open communication on the testing process and/or content. This walk-through of the Practice Test provides opportunities for test administrators to describe the limitations to the support they will be able to provide and the verbal interactions they can have with students during each live test session.

Appendix D. Universal Tools, Designated Supports, and Accommodations

Table 11 –Table 16 outline the embedded and non-embedded universal tools, designated supports, and accommodations that are available to students taking the Montana Science Assessment. See the *Usability, Accessibility, and Accommodations Guidelines* for complete information on student accommodations and supports. See the *TIDE User Guide* for information on how to adjust and set student test settings.

Table 11: Embedded Universal Tools

Universal Tool	Description
Breaks	The number of items per session can be flexibly defined based on the student's need. Breaks of more than 20 minutes will prevent the student from returning to items already attempted by the student. There is no limit on the number of breaks that a student might be given. The use of this universal tool may result in the student needing additional overall time to complete the assessment.
Calculator *(See Non-embedded accommodations for students who cannot use the embedded calculator)*	A grade-appropriate embedded on-screen digital calculator can be accessed for calculator-allowed items when students click on the calculator button. When the embedded calculator, as presented for all students, is not appropriate for a student (for example, for a student who is blind), the student may use the calculator offered with assistive technology devices (such as a talking calculator or a braille calculator). Depending on the student's grade, the following calculator types are provided: • Grade 5- basic calculator • Grade 8- scientific calculator
Digital notepad	This tool is used for making notes about an item. The digital notepad is item-specific and is available through the end of the test segment. Notes are not saved when the student moves on to the next segment or after a break of more than 20 minutes.
Expandable passages and items	Each passage/stimulus and item can be expanded so that it takes up a larger portion of the screen.
Highlighter	A digital tool for marking desired text, item questions, item answers, or parts of these with a color. Highlighted text remains available throughout each test segment.
Keyboard navigation	Navigation throughout text can be accomplished by using a keyboard.
Line reader	The student uses an onscreen universal tool to assist in reading by raising and lowering the tool for each line of text on the screen.
Mark for review	Allows students to flag items for future review during the assessment. Markings are not saved when the student moves on to the next segment or after a break of more than 20 minutes.
Periodic Table	An arrangement of the chemical elements, ordered by their atomic number, electron configuration, and recurring chemical properties. The ordering shows periodic trends, such as elements with similar behavior in the same column.
Strikethrough	Allows users to cross out answer options. If an answer option is an image, a strikethrough line will not appear, but the image will be grayed out.

Table 12: Non-Embedded Universal Tools

Tool	Description
Calculator	When the embedded Desmos Calculator is not suitable for a student participating in the MSA, the provision of a battery-operated hand-held calculator may be appropriate. If a calculator is provided, proctors must ensure that the device is functional, has working batteries, and that the student is familiar and comfortable with how to use it. Students may not use calculators available on their phones, iPads, or other electronic devices. Students may not share calculators. Depending on the student's grade, the following calculator types are permissible: • Grade 5- basic calculator • Grade 8- scientific calculator
English Dictionary	An English dictionary can be provided. The use of this universal tool may result in the student needing additional overall time to complete the assessment.
Scratch paper	Scratch paper to make notes, write computations, or record responses may be made available. A whiteboard with marker may be used as scratch paper. As long as the construct being measured is not impacted, assistive technology devices, including low-tech assistive technology, are permitted to make notes, including the use of digital graph paper. The assistive technology device needs to be familiar to the student and/or consistent with the child's IEP or 504 plan. Access to internet must be disabled on assistive technology devices.
Thesaurus	A thesaurus contains synonyms of terms while a student interacts with text included in the assessment. A full write is the second part of a performance task. The use of this universal tool may result in the student needing additional overall time to complete the assessment.

Table 13: Embedded Designated Supports

Designated Support	Description	Recommendations for Use
Color choices	Enable students to adjust screen background or font color, based on student needs or preferences. This may include reversing the colors for the entire interface or choosing the color of font and background. Yellow on Blue; Medium Gray on Light Gray; Reverse Contrast and Black on Rose	Students with attention difficulties may need this support for viewing test content. It also may be needed by some students with visual impairments or other print disabilities (including learning disabilities). Choice of colors should be informed by evidence that color selections meet the student's needs.
Masking	Masking involves blocking off content that is not of immediate need or that may be distracting to the student. Students can focus their attention on a specific part of a test item by masking.	Students with attention difficulties may need to mask content not of immediate need or that may be distracting during the assessment. This support also may be needed by students with print disabilities (including learning disabilities) or visual impairments. Masking allows students to hide and reveal individual answer options, as well as all navigational buttons and menus.

Table 13: Embedded Designated Supports

Designated Support	Description	Recommendations for Use
Mouse pointer (Size and Color)	This embedded support allows the mouse pointer to be set to a larger size and for the color to be changed. A test administrator sets the size and color of the mouse pointer prior to testing.	Students who are visually impaired and need additional enlargement or a mouse pointer in a different color to more readily find their mouse pointer on the screen will benefit from the mouse pointer support. Students who have visual perception challenges will also find this beneficial. The size and color are set during registration and cannot be changed during the administration of the assessment. Students should have ample opportunity to practice during daily instruction with the size and color to determine student preference. The mouse pointer can be used with the zoom universal tool. If students are using a magnification program (See Designated Support, magnification), the enlarged mouse pointer is built into magnification programs and mouse pointer may not be needed.
Streamlined Mode	This designated support provides a streamlined interface of the test in an alternate, simplified format in which the items are displayed below the stimuli.	This designated support may benefit a small number of students who have specific learning and/or reading disabilities and/or visual impairment in which the text is presented in a more sequential format. Students should have familiarity interacting with items in streamline format.
Text-to-speech	Text is read aloud to the student via embedded text-to-speech technology. The student can control the speed as well as raise or lower the volume of the voice via a volume control.	Students who are struggling readers may need assistance accessing the assessment by having all or portions of the assessment read aloud. This support also may be needed by students with reading-related disabilities, or by students who are blind and do not yet have adequate braille skills. This support will likely be confusing and may impede the performance of students who do not regularly have the support during instruction. Students who use text-to-speech will need headphones unless tested individually in a separate setting.
Translations (toggle)	This Spanish presentation will allow the literate Spanish-speaking student to toggle between a full Spanish translation of the item and the English version of the item. By default, all test directions, navigation buttons, and test content will be presented to the student in the Spanish language.	For students whose primary language is not English and who use dual language supports in the classroom, use of the Spanish translation toggle may be appropriate. Students participate in the assessment regardless of the language. This support will increase reading load and cognitive load. The use of this support may result in the student needing additional overall time to complete the assessment.

Table 13: Embedded Designated Supports

Designated Support	Description	Recommendations for Use
Turn off any universal tools	Disabling any universal tools that might be distracting or that students do not need to use or are unable to use.	Students who are easily distracted (whether or not designated as having attention difficulties or disabilities) may be overwhelmed by some of the universal tools. Knowing which specific tools may be distracting is important for determining which tools to turn off.
Zoom	A tool for making text or other graphics in a window or frame appear larger on the screen. The default font size for all tests is 14 pt. The student can make text and graphics larger by clicking the *Zoom In* button. The student can click the *Zoom Out* button to return to the default or smaller print size. When using the zoom feature, the student only changes the size of text and graphics on the current screen. To increase the default print size of the entire test, the print size must be set for the student in TIDE or set by the test administrator prior to the start of the test. This is the only feature that test administrators can set. The use of this universal tool may result in the student needing additional overall time to complete the assessment.	For students with a visual impairment that require viewing the test in an enlarged print size.

Table 14: Non-Embedded Designated Supports

Support	Description	Recommendations for Use
Amplification	The student adjusts the volume control beyond the computer's built in settings using headphones or other non-embedded devices.	Students may use amplification assistive technology (e.g., headphones, FM System, noise buffers, white noise machines) to increase the volume provided in the assessment platform. Use of this resource likely requires a separate setting. If the device has additional features that may compromise the validity of the test (e.g., internet access), the additional functionality must be deactivated to maintain test security.

Table 14: Non-Embedded Designated Supports

Support	Description	Recommendations for Use
Bilingual dictionary	A bilingual/dual language word-to-word dictionary is a language support. A bilingual/dual language word-to-word dictionary can be provided.	For students whose primary language is not English and who use dual language supports in the classroom, use of a bilingual/dual language word-to-word dictionary may be appropriate. Students participate in the assessment regardless of the language. The use of this support may result in the student needing additional overall time to complete the assessment.
Color contrast	Test content of online items may be printed with different colors.	Students with attention difficulties may need this support for viewing the test when digitally provided color contrasts do not meet their needs. Some students with visual impairments or other print disabilities (including learning disabilities) also may need this support. Choice of colors should be informed by evidence of those colors that meet the student's needs.
Color overlays	Color transparencies are placed over a paper-based assessment.	Students with attention difficulties may need this support to view test content. This support also may be needed by some students with visual impairments or other print disabilities (including learning disabilities). Choice of color should be informed by evidence of those colors that meet the student's needs.
Magnification	The size of specific areas of the screen (e.g., text, formulas, tables, graphics, navigation buttons, and mouse pointer) may be adjusted by the student with an assistive technology device or software. Magnification allows increasing the size and changing of the color contrast, including the size and color of the mouse pointer, to a level not provided for by the zoom universal tool, color contrast designated support, and/or mouse pointer designated support.	Students used to viewing enlarged text or graphics, or navigation buttons with or without changes to color contrast, may need magnification to comfortably view content. This support also may meet the needs of students with visual impairments and other print disabilities. The use of this designated support may result in the student needing additional overall time to complete the assessment.
Medical supports	Students may have access to medical supports for medical purposes (e.g., Glucose Monitor). The medical support may include a cell phone and should only support the student during testing for medical reasons.	Educators should follow local policies regarding medical supports and ensure students' health is the highest priority. Electronic medical support settings must restrict access to other applications, or the test administrator must closely monitor the use of the medical support to maintain test security. Use of medical supports may require a separate setting to avoid distractions to other test takers and to ensure test security.

Table 14: Non-Embedded Designated Supports

Support	Description	Recommendations for Use
Noise buffers	Ear mufflers, white noise, and/or other equipment used to block external sounds.	Student (not groups of students) wears equipment to reduce environmental noises. Students may have these testing variations if regularly used in the classroom. Students who use noise buffers will need headphones unless tested individually in a separate setting.
Read aloud	Text is read aloud to the student by a trained and qualified human reader who follows the administration guidelines provided in the *Test Administration Manual* and *Read Aloud Guidelines* (see Montana Usability, Accessibility, and Accommodations Guidelines). All or portions of the content may be read aloud.	Students who are struggling readers may need assistance accessing the assessment by having all or portions of the assessment read aloud. This support also may be needed by students with reading-related disabilities, or by students who are blind and do not yet have adequate braille skills. If not used regularly during instruction, this support is likely to be confusing and may impede the performance on assessments. Readers should be provided to students on an individual basis – not to a group of students. A student should have the option of asking a reader to slow down or repeat text. The use of this support may result in the student needing additional overall time to complete the assessment and/or the use of a separate setting.
Read aloud in Spanish	Spanish text is read aloud to the student by a trained and qualified human reader who follows the administration guidelines provided in the *Test Administration Manual* and the *Read Aloud guidelines*. All or portions of the content may be read aloud.	Students receiving the translations (stacked) designated support and who are struggling readers may need assistance accessing the assessment by having all or portions of the assessment read aloud. This support also may be needed by students with reading-related disabilities. If not used regularly during instruction, this support is likely to be confusing and may impede the performance on assessments. A student should have the option of asking a reader to slow down or repeat text. The use of this support may result in the student needing additional overall time to complete the assessment and/or the use of a separate setting.
Scribe	Students dictate their responses to a human who records verbatim what they dictate. The scribe must be trained and qualified and must follow the administration guidelines provided in the *Test Administration Manual.*	Students who have documented significant motor or processing difficulties, or who have had a recent injury (such as a broken hand or arm) that make it difficult to produce responses may need to dictate their responses to a human, who then records the students' responses verbatim. The use of this support may result in the student needing additional overall time to complete the assessment.

Table 14: Non-Embedded Designated Supports

Support	Description	Recommendations for Use
Separate setting	Test location is altered so that the student is tested in a setting different from that made available for most students.	Students who are easily distracted (or may distract others) in the presence of other students, for example, may need an alternate location to be able to take the assessment. The separate setting may be in a different room that allows them to work individually or among a smaller group. The student may read aloud to self, use a device requiring voicing (e.g., a Whisper Phone), or use Amplification. It may also include a calming device or support as recommended by educators and/or specialists. Or, the separate setting may be in the same room but in a specific location (for example, away from windows, doors, or pencil sharpeners, in a study carrel, near the teacher's desk, or in the front of a classroom). Some students may benefit from being in an environment that allows for movement, such as being able to walk around. In some instances, students may need to interact with instructional or test content outside of school, such as in a hospital or their home. A specific adult, trained in a manner consistent with the TAM, can act as test proctor (test administrator) when student requires it.
Simplified Test Directions	The test administrator simplifies or paraphrases the test directions found in the *Summative Test Administration Manual.*	Students who need additional support understanding the test direction may benefit from this resource. This designated support may require testing in a separate setting to avoid distracting other test takers.
Translated Test Directions	PDF of directions translated in each of the languages currently supported. Bilingual adult can read to student.	Students who have limited English language skills (whether or not designated as ELs or ELs with disabilities) can use the translated test directions. In addition, a biliterate adult trained in the test administration manual can read the test directions to the student. The use of this support may result in the student needing additional overall time to complete the assessment.

Table 15: Embedded Accommodations

Accommodation	Description	Recommendations for Use
Braille	Braille is available in the following presentation: • UEB Contracted with Nemeth Math	Students with visual impairments may read text via braille. Tactile overlays and graphics also may be used to assist the student in accessing content through touch. Alternative text descriptions are embedded in the assessment for all graphics. The type of braille presented to the student (contracted or non-contracted) is set in TIDE. The use of this accommodation may result in the student needing additional overall time to complete the assessment.

Table 16: Non-Embedded Accommodations

Accommodation	Description	Recommendations for Use
Alternate response options	Alternate response options include but are not limited to adapted keyboards, large keyboards, Sticky Keys, Mouse Keys, FilterKeys, adapted mouse, touch screen, head wand, and switches.	Students with some physical disabilities (including both fine motor and gross motor skills) may need to use the alternate response options accommodation. Some alternate response options are external devices that must be plugged in and be compatible with the assessment delivery platform.
Braille (paper/pencil assessment form)	A raised-dot code that individuals read with the fingertips. Graphic material (e.g., maps, charts, graphs, diagrams, and illustrations) is presented in a raised format (paper or thermoform).	Students with visual impairments may read text via braille. Tactile overlays and graphics also may be used to assist the student in accessing content through touch. The use of this accommodation may result in the student needing additional overall time to complete the assessment.
Print on demand	Paper copies of either passages/stimuli and/or items are printed for students. For those students needing a paper copy of a passage or stimulus, permission for the students to request printing must first be set in TIDE. For those students needing a paper copy of one or more items, the help desk must be contacted by the school or district coordinator to have the accommodation set for the student.	Some students with disabilities may need paper copies of either passages/stimuli and/or items. A very small percentage of students should need this accommodation. The use of this accommodation may result in the student needing additional time to complete the assessment.

Accommodation	Description	Recommendations for Use
Sign Language for Test Items	Test content is translated by an ASL human signer.	Some students who are deaf or hard of hearing and who typically use ASL may need this accommodation when accessing text-based content in the assessment. The use of this accommodation may result in the student needing additional overall time to complete the assessment. For many students who are deaf or hard of hearing, viewing signs is the only way to access information presented orally. It is important to note, however, that there is no required audio as part of the Montana Science Assessment.
Specialized Calculator	A non-embedded calculator for students needing a special calculator, such as a braille calculator or a talking calculator, currently unavailable within the assessment platform.	Students with visual impairments who are unable to use the embedded calculator for calculator-allowed items will be able to use the calculator that they typically use, such as a braille calculator or a talking calculator. Test administrators should ensure that the calculator is available only for designated calculator items.

Accommodation	Description	Recommendations for Use
Speech-to-text	Voice recognition allows students to use their voices as input devices to the computer, to dictate responses or give commands (e.g., opening application programs, pulling down menus, and saving work). Voice recognition software generally can recognize speech up to 160 words per minute. Students may use their own assistive technology devices.	Students who have motor or processing disabilities (such as dyslexia) or who have had a recent injury (such as a broken hand or arm) that make it difficult to produce text or commands using computer keys may need alternative ways to work with computers. Students will need to be familiar with the software, and have had many opportunities to use it prior to testing. Speech-to-text software requires that the student go back through all generated text to correct errors in transcription, including use of writing conventions; thus, prior experience with this accommodation is essential. If students use their own assistive technology devices, all assessment content should be deleted from these devices after the test for security purposes. For many of these students, using voice recognition software is the only way to demonstrate their composition skills. Still, use of speech-to-text does require that students know writing conventions and that they have the review and editing skills required of students who enter text via the computer keyboard. It is important that students who use speech-to-text also be able to develop planning notes via speech-to-text, and to view what they produce while composing via speech-to-text.
Timing or Scheduling	Students may require special timing or scheduling considerations.	Students who require special or unique timing or scheduling considerations, where they may not be able to test with other students or the rest of their class, etc.

Accommodation	Description	Recommendations for Use
Word prediction	Word prediction allows students to begin writing a word and choose from a list of words that have been predicted from word frequency and syntax rules. Word prediction is delivered via a non-embedded software program. The program must use only single word prediction. Functionality such as phrase prediction, predict ahead, or next word must be deactivated. The program must have settings that allow only a basic dictionary. Expanded dictionaries, such as topic dictionaries and word banks, must be deactivated. Phonetic spelling functionality may be used, as well as speech output built into the program which reads back the information the student has written. If further supports are needed for speech output, see text-to-speech or read aloud policies. Students who use word prediction in conjunction with speech output will need headphones unless tested individually in a separate setting. Students may use their own assistive technology devices.	Students who have documented motor or orthopedic impairments, which severely impairs their ability to provide written or typed responses without the use of assistive technology, may use word prediction. Students with moderate to severe learning disabilities that prevent them from recalling, processing, or expressing written language may also use word prediction. Students will need to be familiar with the software, and have had many opportunities to use it in daily instruction. Use of word prediction does require that students know writing conventions and that they have the review and editing skills required of all students. It is important that students who use word prediction also be able to develop planning notes and review their writing with or without text-to-speech. If students use their own assistive technology devices, all assessment content should be deleted from these devices after the test for security purposes.

Appendix E. Examples of Test Security Violations

Table 17: Test Security Violations

Term	Level I	Level II	Level III
Description	(Unusual circumstances or routine testing phenomena)	(Non-routine or unusual circumstances)	(Rare and extreme events)
Reporting Timeframe	Reported to the BC within 5 days. (Unless otherwise specified in the assessment specific TAM.)	BC reports to the OPI Assessment Division in the MontCAS Application within 3 days. (Unless otherwise specified in the assessment specific TAM.)	Reported to the OPI Assessment Division immediately by phone (1.844.867.2569). BC must report in the MontCAS Application within 24 hours. (Unless otherwise specified in the assessment specific TAM.)
Program Impact	**Low**	**Medium**	**High**
	Low impact on individual or group of testing students	May impact the performance of a student or group of testing students	May impact large group of testing students, state or nation and compromise the test
Risks	Little risk to interpretation of test scores	Has the potential to skew the interpretation of test scores	Severe risk to test program or interpretation of test scores
Features of Incident	• Low risk of impacting: ○ Student performance ○ Test security ○ Test validity	• Has little or no impact on validity of the assessment • No statewide or national implications • May result in invalidation of test results	• May expose secure test materials • May affect all students participating - statewide and nationwide

The OPI's definition of each Level:

- Level I – This includes many different activities, not necessarily cheating, but anything unusual that happened during testing. The event is considered routine, but the standards of test administration were not observed. It has a low impact on the individual or group of students who are testing and has a low risk of potentially affecting student performance on the test, test security, or test validity.
    - **Examples**: Student(s) leave the test room without authorization or student(s) making distracting gestures/sounds or talking during the test session that creates a disruption in the test session for other students.

- Level II – This includes many different activities; however, it is more serious than a Level I and the standards of test administration were not observed. The difference between Level I and Level II is usually defined in perception of the degree, intent, and/or effect of the misconduct. It may impact the performance of a student or group of students who are testing and may potentially affect student performance on the test or interpretation of those scores. (Typically, does not have statewide or nationwide implications).

    - Examples: Administrator left instructional related materials on the walls or desks in testing environment or administrator allowed inappropriate designated supports and/or accommodations during test window.

- Level III – This is a narrow set of activities; it is the most serious Level as it describes the act of breaking or failing to observe the standards of test administration. The difference between Level II and Level III is defined in perception of the intent and/or effect of the misconduct. It may compromise the integrity of an assessment, typically by exposing secure test materials. The implications of Level III can affect all students participating in the assessment statewide and nationwide.
    - Examples: Administrator or coordinator modified a student's response or records, or an adult or student posted item(s) or test materials on social media, website, or any other form of media.

Appendix F. System Test Coordinator Checklist

The System (District) Test Coordinator activity checklist list below provides recommended steps to help test coordinators prepare for an online test administration.

	District Test Coordinator Activities	Manual in reference	Estimated Time to Complete	Target Completion Date	Notes/Resources
	Direct Responsibilities				
☐	Notify schools when administration manuals and training materials are available.	*TAM*, Section II		Complete at least 5–6 weeks prior to testing.	https://mt.portal.cambiumast.com/
☐	Review scheduling and testing requirements with BCs.	*TAM*, Section VI	60–90 minutes	Complete at least 3–4 weeks prior to testing.	Section VI
☐	Plan any state-required training for BCs and test administrators; retrain as needed.			Complete at least 3 weeks prior to testing.	
☐	Review training modules and schedule training session for SCs.	*TAM*, Section II	2–3 hours	Complete while spaced over 2–3 weeks prior to testing.	Section II
☐	Review security procedures with BCs and test administrators.	*TAM*, Section III, *Appendix E*		Complete at least 2–3 weeks prior to testing.	Section III
☐	Work with District Technology Coordinators to ensure timely computer setup. • Conduct network diagnostics. • Download the secure browser (beginning late August or early September). • Verify that schools meet the minimum technology requirements.		5–10 hours	Complete at least 2 weeks prior to testing.	*Technology Guide*
☐	Work with schools to review TIDE student enrollment information.		2–4 hours	Complete at least 2 weeks before testing and during testing.	*TIDE User Guide*
☐	Perform an equipment needs assessment based on individual student requirements. • Work with the SC to identify students who will need specialized		1–2 hours	Complete at least 2 weeks prior to testing.	Students are permitted and encouraged to use their own ear buds or headsets—but districts and schools should also plan on having some available.

District Test Coordinator Activities	Manual in reference	Estimated Time to Complete	Target Completion Date	Notes/Resources
equipment for accommodations.				
Monitor Administration Activities				
☐ Investigate all testing improprieties, irregularities, and breaches reported by test administrators. • Follow reporting procedure according to Section IV.	*TAM, Appendix E*	As needed	Ongoing. Deadline for all submissions will be one week after the testing window closes.	Section IV
☐ Report any test security incident in TIDE and report to the appropriate personnel according to the guidelines.	*TAM, Appendix E*	As needed		Section IV
☐ Monitor with the District Technology Coordinator any technical problems to apply resolutions or disseminate information at other school sites prior to testing.			Ongoing throughout the testing window.	
Oversight Responsibilities				
Ensure that BCs and test administrators in the district are appropriately trained and aware of policies and procedures, especially related to security.	*TAM, Appendix E*		Complete at least 3 weeks prior to testing	
Review school test administration schedules for adequate time and resource planning.	*TAM*, Section VI.III		Complete at least 3 weeks prior to testing	
Verify that BCs and test administrators have reviewed student information in TIDE and are verifying student settings for designated supports and accommodations in TIDE.			Complete at least 2 weeks prior to testing.	*TIDE User Guide*
You or your designee must be available during your testing window for questions and problem solving.		Complete during your entire testing window.		
Communicate regularly with BCs any emerging trends or issues.		Complete during your entire testing window.		

Appendix G. School Test Coordinator Checklist

The School Test Coordinator activity checklist list below provides recommended steps for a coordinator to help prepare for an online test administration.

	School Test Coordinator Activities	Reference in Manual	Estimated Time to Complete	Target Completion Date	Notes/Resources
	Direct Responsibilities				
☐	Attend trainings your district offers and review all Montana Science policy and test administration documents.		60–90 minutes	Complete at least 3 weeks prior to testing.	https://mt.portal.cambiumast.com/
☐	Work with technology personnel to ensure timely computer setup: • Conduct network diagnostics. • Download the secure browser. • Verify that your school has met the minimum technology requirements. • Ensure that other technical issues are resolved before and during testing.		5–10 hours	Complete at least 3–4 weeks prior to testing.	*Technology Guide*
☐	Perform an equipment needs check based on individual student requirements. • Work with test administrators to identify students who will need specialized equipment for accommodations.		1–2 hours	Complete at least 2 weeks prior to testing.	
☐	Based on the test administration windows, work with test administrators and ARs to establish a testing schedule.	*TAM,* Section VI.III	2–4 hours	Complete at least 1–2 weeks prior to testing.	
☐	Work with test administrators to review student information in TIDE before students are tested to ensure that correct student information and test settings for designated supports and accommodations are applied.		2–4 hours	Complete at least 1–2 weeks prior to testing.	*TIDE User Guide*

School Test Coordinator Activities	Reference in Manual	Estimated Time to Complete	Target Completion Date	Notes/Resources
☐ Establish a place to test those students who need a separate test setting.	*TAM,* Section VI.IV	1–2 hours	Complete at least 1–2 weeks prior to testing.	*Usability, Accessibility, and Accommodations Guidelines*
☐ Work with test administrators to plan a quiet activity for each test session for students who finish early.		30 minutes	Complete the week of testing.	The activity should not be related to the test being given. For example, students who finish early may work on assignments for unrelated subjects or read a book.
☐ Ensure proper handling of all printed test materials and scratch paper. Collect all test materials on each day of testing to keep in a secure location until after the test session, and then destroy according to the security policy outlined in the *TAM*.	*TAM,* Section III		Complete during and after testing window.	
☐ Ensure adherence to all security policies. • Ensure that all test administrators read and sign state security agreements if required by your state.	*TAM,* Section III		Ongoing	
☐ Document any testing impropriety, irregularity, or breach and report to the AR immediately after learning of the incident. • Working with the AR, enter incidents in TIDE.	*TAM,* Section IV	As needed	Ongoing	
Monitor Administration Activities				
☐ Monitor testing progress during the testing window and ensure that all students participate as appropriate, addressing student issues as needed.		As needed	Ongoing	
☐ Raise any technical issues with the School Technology Coordinator for resolution.		As needed	Ongoing	
☐ Review, investigate, and report on all potential testing improprieties, irregularities, and breaches reported by the test administrator. Mitigate incidents when appropriate.	*TAM,* Section IV	As needed	Ongoing	

	School Test Coordinator Activities	Reference in Manual	Estimated Time to Complete	Target Completion Date	Notes/Resources
	Oversight Responsibilities				
☐	Verify that test administrators have verified student settings for designated supports and accommodations in TIDE.		As needed	Complete at least 1–2 weeks prior to testing.	*TIDE User Guide*
☐	Assist test administrators in taking proper measures to ensure that students have access only to allowable non-embedded resources during testing.		As needed	Complete prior to testing.	
☐	Verify that test administrators are adhering to all test security policies and practices and that they have access to the Test Security Incident Log, and that they understand how to complete this document if testing improprieties, irregularities, or breaches occur.	*TAM,* Section IV	As needed	Complete during testing.	

Appendix H. Test Administrator Checklist

The Test Administrator activity checklist list below provides recommended steps to help prepare a TA for an online test administration.

	Test Administrator Activities	Reference in Manual	Estimated Time to Complete	Target Completion Date	Notes/Resources
☐	Review all Montana Science policy and test administration documents, particularly the *Test Administration Manual (TAM).*	*TAM*	1–2 hours	Complete at least 2–3 weeks prior to testing.	https://mt.portal.cambiumast.com/
☐	Review the training modules and attend school or district training sessions, if any are offered.	*TAM,* Section I.III	2–3 hours	Complete at least 2–3 weeks prior to testing.	https://mt.portal.cambiumast.com/
☐	Provide students with a walk-through of the Training Test and/or Practice Test for familiarity with navigation of the system and tools	*TAM, Appendix C*	1 hour	Complete while spaced over 2–3 weeks prior to testing.	https://mt.portal.cambiumast.com/science.html
☐	Perform an equipment needs check based on individual student requirements. • Work with the BC to identify students who will need specialized equipment for accommodations.		1–2 hours	Complete at least 2 weeks prior to testing.	Students are permitted and encouraged to use their own ear buds or headsets—but districts and schools should also plan on having some available.
☐	Work with the BC to determine precise testing schedules based on the test administration windows selected by the school schedule. • Make sure your students' test administration schedule includes allowable breaks.	*TAM*, Section VI.III		Complete at least 1–2 weeks prior to testing.	
☐	Confirm that you have received your TIDE login information. You should have received an automated email from the Help Desk notifying you of how to log in to TIDE. You will also use this username			Complete at least 1–2 weeks prior to testing.	If you have not received this information, please check your spam/junk email folder to see if it was mistakenly routed there. If not, check with your School Test Coordinator. *TIDE User Guide*

Test Administrator Activities	Reference in Manual	Estimated Time to Complete	Target Completion Date	Notes/Resources
and password for the Online Testing System.				
☐ Work with your BC to ensure that each student has an SSID number and has been loaded into TIDE.			Complete at least 1–2 weeks prior to testing.	*TIDE User Guide*
☐ Confirm each student's test settings for designated supports and accommodations in TIDE against their IEP or other relevant documentation as appropriate.		2–4 hours	Complete at least 1–2 weeks prior to testing.	*TIDE User Guide*
☐ Ensure that the MT secure browser has been downloaded to any computer(s) on which students will be testing.		1–2 hours	Complete at least 1–2 weeks prior to testing. *and again* Complete the day before testing or morning of testing.	Check with your BC or School Technical Coordinator if the [**Secure Browser**] icon is missing. *Technology Guide*
☐ Plan a quiet activity for each test session for students who finish early.		15–30 minutes	Complete the week of testing.	The activity should not be related to the test being given. For example, students who finish early may work on assignments for unrelated subjects or read a book.
☐ Prior to administration, check all computers that will be used and close all applications except those identified as necessary by the school's Technology Coordinator. Make sure that no computer has dual monitors.		1–2 hours	Complete the morning of testing.	The TA should open the secure browser on each computer after closing any unnecessary applications.
☐ Review all guidelines for creating a secure test environment. • Review all security procedures and guidelines in the *TAM*. • Carefully read and sign a test security agreement if required by your state.	*TAM,* Section VI		Prior to and during day(s) of testing.	

	Test Administrator Activities	Reference in Manual	Estimated Time to Complete	Target Completion Date	Notes/Resources
☐	Make sure the physical conditions of the testing room are satisfactory. • Make sure that no instructional materials directly related to the content of the assessments are visible. • Students should be seated so there is enough space between them, or provide desktop partitions to minimize opportunities to look at each other's screen. • Actively monitor students throughout the test sessions. • Students who are not being tested may not be in the room where a test is being administered.	*TAM,* Section VI		Complete the day(s) of testing.	• Make sure students clear their desks and put away all books, backpacks, purses, cell phones, electronic devices of any kind, and other materials not needed for the test. • Ensure that students do not access unauthorized electronic devices (such as cell phones) at any time during testing.
☐	On the day of testing, verify that the students have their login information (first name, SSID, and session ID).	*TAM,* Section VIII.I		Complete the day(s) of testing.	*TIDE User Guide*
☐	Report any testing improprieties, irregularities, and breaches to the BC and STC in writing immediately following an impropriety, irregularity, or breach.	*TAM,* Section IV		Complete as soon as possible during or immediately following testing.	
☐	Securely dispose of all printed testing materials, including student login information, print-on-demand documents, and scratch paper in a secure manner.	*TAM,* Section III		Complete after testing.	

Assessment Office Contact Information

The contact information for Assessment Division of the Montana Office of Public Instruction.

Contact Information	
Questions about state policies	**Questions about technology and the overall administration procedures**
Montana OPI **Accommodations, Test Policy, Testing Irregularities** Name: OPI Assessment Phone: 844-867-2569 Email: OPIAssessmentHelpDesk@mt.gov	Montana Help Desk Monday–Friday 6:00 a.m. to 6:00 p.m. Mountain Time 1-877-365-7915 mthelpdesk@cambiumassessment.com

Appendix I. Change Log

Change Made	Section	Date

Appendix 5-B
Assistive Technology Manual

Montana Office of Public Instruction

Assistive Technology Manual for Windows & macOS

2022–2023

Published January 30, 2023

Prepared by Cambium Assessment, Inc.

Table of Contents

Overview of Testing with Assistive Technology

This manual provides an overview of the embedded and non-embedded assistive technology tools that can be used to help students with accessibility needs complete online tests in the Test Delivery System (TDS). It includes lists of supported devices and applications for each type of assistive technology that students may need, as well as setup instructions for the assistive technologies that require additional configuration in order to work with TDS.

- Embedded assistive technology tools include the built-in test tools in TDS, such as the text-to-speech and speech-to-text tools. These tools can be accessed without third-party software or hardware and do not require Permissive Mode to be turned on in TDS.
  - The embedded assistive text-to-speech, speech-to-text, and zoom tools in TDS are available on supported desktop and mobile devices. However, the embedded word prediction tool is not available on Linux or iPadOS.
- Non-embedded assistive technology tools are the third-party accessibility hardware and software that students use to help them complete tests in TDS. These tools require Permissive Mode to be turned on in TDS and may require additional configuration steps prior to testing.
  - **Most of the non-embedded assistive technology tools covered in this manual are supported by Windows and macOS desktops only.** Mobile devices, such as Chromebooks and iPads, do not support Permissive Mode, which is required for non-embedded technology. More detailed system requirements are provided for each assistive technology tool throughout this manual.
  - Permissive Mode on macOS is supported only with Secure Browser versions 12.5 and 15.5. Ensure that the proper Secure Browser version is downloaded from the Montana Testing Portal to allow for testing with Permissive Mode.

Students who use assistive technology tools to interact with a standard web browser should be able to use those same tools with TDS, unless they are web-based applications or browser extensions. The best way to test compatibility for tools is to take a practice test in the Secure Browser with those tools turned on. If they do not work, see the additional configuration instructions in this manual as required. If you still have questions about the assistive technology covered in this guide, please contact the Help Desk.

For more information about taking practice tests, see the *Test Administrator User Guide*.

The manual includes the following sections:

- Testing with Speech-to-Text Technology
- Testing with Word Prediction Technology
- Testing with Alternative Computer Input Technology
- Testing with Assistive Keyboard and Mouse Input Technology
- Testing with Screen Magnifier Technology
- Testing with Text-to-Speech
- Testing with Assistive Technology for Braille Tests

Using Permissive Mode with Assistive Technology

Permissive Mode is a TDS accommodation that allows students to use non-embedded assistive technology to complete tests in the Secure Browser. It must be turned on for any students testing with third-party assistive technology tools. When Permissive Mode is turned on, the Secure Browser's security settings will be partially lowered to allow students to use tools that would otherwise be blocked. This accommodation should be assigned to students in TIDE before they begin testing. For more information, see the *TIDE User Guide*.

Permissive Mode is available for computers running supported desktop Windows and macOS operating systems only. Permissive Mode on macOS devices is supported only with Secure Browser versions 12.5 and 15.5. When using Windows devices, the task bar remains on-screen throughout the test after enabling accessibility software. However, forbidden applications are still prohibited.

Permissive Mode activates when students are approved for testing in TDS. The student's assistive technology should already be set up for use with TDS when they begin testing with Permissive Mode.

When Permissive Mode is turned on, standard keyboard commands in the Secure Browser will be disabled in order to accommodate any potential keyboard commands associated with the assistive technology the student may be using. For information about standard keyboard commands in the Secure Browser, see the *Test Administrator User Guide*.

Note: Some of the tools covered in this document may include prohibited features that students can access, even when Permissive Mode is enabled. For such tools, it is recommended that Test Administrators monitor students while testing to ensure they do not use any prohibited features.

How to Use Assistive Technology with Permissive Mode on Windows and macOS 10.15

Students can follow the instructions in this section to use assistive technology alongside the Secure Browser with all supported Windows operating systems and with macOS 10.15.

1. Open the required accessibility software on the student’s testing device.
2. Open the Secure Browser. Begin the standard sign-in process up to the proctor approval step.
3. When a student is approved for testing, the Secure Browser allows the operating system’s menu and task bar to appear.
   - **Windows:** The Secure Browser resizes, and the taskbar remains visible inside the test in its usual position. Students can press **Alt**+**Tab** to switch between the Secure Browser and accessibility applications that they are permitted to use in their test session.
   - **macOS 10.15:** The Secure Browser resizes, and students can view the dock in its usual position inside the test. If the dock is set to autohide, no resizing occurs, and the dock is visible only when the mouse moves toward the bottom of screen. Students can press **Cmd**+**Tab** to switch between the Secure Browser and permitted accessibility applications.

4. The student must immediately switch to the accessibility software that is already open on the computer so that it appears over the Secure Browser. The student cannot click within the Secure Browser until the accessibility software is configured.

   – **Windows:** Click the accessibility software application in the task bar.

   – **macOS 10.15:** Click the accessibility software application in the dock.

5. The student configures the accessibility software settings as needed.

   – If the student needs to test with multiple accessibility applications simultaneously, they should configure all of them during this step.

6. After configuring all the necessary the accessibility software settings, the student returns to the Secure Browser and continues the sign-in process. At this point, the student can no longer switch back to the accessibility software. If changes need to be made to the accessibility software, the student must sign out and then sign in again.

How to Use Assistive Technology with Permissive Mode on macOS 13

The Secure Browser 15.5 for macOS 13 requires students to launch approved accessibility software after logging in to the test.

1. Open the required accessibility software on the student's testing device and configure the accessibility software settings as needed.

2. Open the Secure Browser and follow the standard sign-in process.

3. On the first page of the test, click **Applications** in the global menu. From the ***App Launcher*** window that pops up, select the appropriate assistive technology software and click **OK**.

4. Return to the Secure Browser and continue testing.

   – **Note**: Students should not close the assistive technology application while the Secure Browser is still open, as they may not be able to relaunch it. If the student accidentally closes the assistive technology application, they must sign out and then sign in again.

Testing with Speech-to-Text Technology

Speech-to-text (STT) technology transcribes a student's spoken words into text for item responses in TDS. Students with the appropriate accommodations may use STT assistive technology while taking tests. It is recommended that students use the embedded STT tool in TDS. If a third-party STT tool still needs to be used, please contact the Assessment Help Desk at the Montana OPI at 1-844-867-2569, as state approval will be required prior to be able to download any third-party software.

Students using STT technology will need to use a headset while testing. Any wired headset with a 3.5mm or USB connection should be compatible.

Testing with Word Prediction Technology

Word prediction assistive technology suggests words to students as they type responses for test items. TDS does not include any embedded predictive text tools, but it supports several third-party toolst aht use predictive text technology.

Table 1 provides the technology requirements for students testing with word prediction assistive technology. Please note that if students need to use the **Read&Write for Windows & Mac** application, Test Administrators should reach out to the Assessment Help Desk at the Montana OPI at 1-844-867-2569, as state approval will be required prior to being able to download.

Table 1. Third-Party Word Prediction Applications

Product	System Requirements	Additional Details
WordQ+SpeakQ (WordQ5) • Supported Version: 5.3.5	• Windows 8.1, 10; Server 2012 R2, 2016 R2	• Requires additional setup before use in TDS (see configuration information) ▪ Exam Mode must be enabled before students begin testing.
Read&Write for Windows & macOS • Supported macOS Versions: 7.1.3, 7.1.5, 7.1.6 • Supported Windows Version: 12.0.73, 12.0.75	• R&W 7.1.3 is compatible with macOS 10.15 • R&W 7.1.5 is compatible with macOS 12 • R&W 7.1.6 is compatible with macOS 13 • R&W 12.0.75 is compatible with Windows 11 • R&W 12.0.73 is compatible with Windows 8.1, 10; Server 2012 R2, 2016 R2	• TDS cannot confirm appropriate configurations are in use during an exam, so students may be able to access prohibited features. • Requires additional setup before use in TDS (see configuration information) • The Windows version also includes speech-to-text functionality that students may use if they have the proper accommodations. • To use text prediction feature in Windows, students can press the available function keys from F1 to F8. • Includes a Screen Recording feature that should be disabled for Read and Write via **System Settings** > **Privacy & Security** > **Screen Recording**.

WordQ5

To minimize security risks, WordQ includes an Exam Mode feature, which can be enabled through the application's settings. Exam Mode requires a time limit of 1–12 hours to be set. Please note, this does not eliminate all security risks, and once Exam Mode has been set, it cannot be disabled until the configured time has run out.

To turn on exam mode, click the **Options** icon, and select **Exam Mode**. In the dialog popup window that appears, you can allow and restrict the **Word usage examples** and **Single words added by the user including topic words** features. You can also set the exam time limit at the bottom of the window.

Read&Write (Windows & macOS)

Read&Write has an Exam Mode that can be used to turn off features for a single student on their particular testing device. When Exam Mode is enabled, the student will have access to only the selected features on the toolbar. For students requiring word prediction, Word Prediction should be enabled. Read&Write for Windows also includes text-to-speech and speech-to-text (talk & type) features, which should be enabled for students with the proper accommodation settings only.

To use Exam mode, run Read & Write and click the settings button in the top-right corner, then click **Show more settings**. In the *Find a Setting* field, type *adminsettings*. Enter and confirm a password to grant access on this computer. This username and password are associated with the administrative account for your Read&Write subscription. When logged in to administrator settings, click the **Select your features** tab and select which features you'd like to be enabled on the student's toolbar. Enable the **Use Exam Mode now** toggle to start Exam Mode, then close the Read&Write menu to start the exam.

Testing with Alternative Computer Input Technology

Alternative Computer Input (ACI) assistive tools allow students with various impairments (such as physical and visual impairments) to interact with a computer without using a traditional mouse and keyboard setup. For instance, Dwell Clicker 2 allows students to use a mouse without having to click the left or right mouse buttons.

TDS does not include any embedded alternative computer input tools, but it supports several third-party alternative computer input technologies.

Table 2 provides a list of third-party ACI devices that can be used in TDS. Please note that this list includes only the devices that CAI has thoroughly tested against the Secure Browser, but there may be additional supported ACI devices that have not been tested yet. If your students need to use an ACI device not listed here, please test it in a practice test first to ensure there are no issues with it.

Table 2. Third-Party ACI Devices

Product	System Requirements	Additional Details
Dwell Clicker 2 • Supported Versions: 2.0.1.8	• Windows 10, 11; Server 2012 R2, 2016 R2	• Requires additional setup before use in TDS (see configuration instructions)
HeadMouse Nano	• Windows 8.1, 10, 11; Server 2012 R2, 2016 R2 • macOS 10.15 • macOS 11.6 • macOS 12.4	• Requires additional setup before use in TDS (see configuration instructions)
Access Switch	• Windows 8.1, 10, 11; Server 2012 R2, 2016 R2 • macOS 10.15 • macOS 11.6 • macOS 12.4 • macOS 13	N/A
Swifty • Supported Versions: SW2	• Windows 8.1, 10, 11; Server 2012 R2, 2016 R2 • macOS 10.15 • macOS 11.6 • macOS 12.4 • macOS 13	• Requires additional setup before use in TDS (see configuration instructions)

Configuring Dwell Clicker 2

To configure Dwell Clicker 2 settings, open the application and select the keyboard icon, then click the **Options** key. In the window that pops up, make sure the **Use Text Prediction** checkbox is not checked.

Configuring HeadMouse Nano

To configure HeadMouse Nano when using the SofType keyboard, open the SofType application and select **View** > **Word Bar** from the menu. Then make sure the **Prediction** radio button is not marked.

Configuring HeadMouse Nano for macOS

The HeadMouse Nano for OSX can be used to mimic mouse-clicking movements only in conjunction with an Access Switch device (such as an AbleNet Switch) and the regular Apple on-screen keyboard. When completing a test with a Switch, students can left-click, drag-and-drop, double-click, and right-click (right-clicking would require an additional Switch).

To configure HeadMouse Nano when using the Apple on-screen keyboard, open **System Preferences** > **Keyboard** > **Text**. Then make sure the following checkboxes are unchecked:

- **Add period with double-space**
- **Capitalize words automatically**
- **Correct spelling automatically**

Configuring Swifty: SW2

To configure Swifty Switch Access for the student's needs, the following DIP Switches should be set. After you modify DIP Switch settings, unplug and re-plug Swifty to activate the settings.

Table 3. DIP Switches

Switch 1	Switch 2	USB Device	Interface Actions
ON	ON	Mouse	Left, Right, Middle
OFF	ON	Joystick	Btn1, Btn2, Btn3
ON	OFF	Keyboard (For iPad)	Enter, Space, Tab
ON	OFF	Keyboard	1,2, 3

Testing with Assistive Keyboard and Mouse Input Technology

Assistive Keyboard and Mouse Input tools provide additional support to students with physical impairments who need to use a keyboard and mouse in order to respond to test items. These include keyboards with larger keys, computer mice with trackballs, and other tools that make it easier for students with limited movement abilities to use a computer.

TDS does not include any embedded assistive keyboard and mouse input tools, as these tools typically involve the use of special hardware, but TDS does support several third-party assistive keyboard and mouse input tools.

Table 4 provides a list of third-party assistive keyboard and mouse input tools that can be used in TDS. Please note, there may be additional supported assistive keyboards and mouse input tools that have not been tested yet. If your students need to use a device not listed here, please test it in a practice test first to ensure there are no issues.

Some third-party assistive keyboards have special function keys that put the computer to sleep. If pressed, the computer will go to sleep, and the student will be kicked out of the test and have to sign back in to resume testing.

Table 4. Third-Party Assistive Keyboard and Mouse Input Technology

Product	System Requirements	Additional Details
Keys-U-See Keyboard	• Windows 10, 11; Server 2012 R2, 2016 R2	N/A
BigKeys Keyboard • Supported Versions: Plus, LX	• Windows 8.1, 10, 11; Server 2012 R2, 2016 R2 • macOS 10.15 • macOS 11.6 • macOS 12.4 • macOS 13	N/A
BigTrack2 Trackball	• Windows 8.1, 10, 11; Server 2012 R2, 2016 R2 • macOS 10.15 • macOS 11.6 • macOS 12.4 • macOS 13	N/A

Testing with Screen Magnifier Technology

Screen magnifier assistive technology enlarges the content displayed on the computer screen in order to assist students with visual impairments. Although TDS supports some non-embedded screen magnifier tools from third parties, it is recommended that students use the embedded zoom tools in TDS. These embedded tools were designed to magnify test content in the most intuitive and user-friendly manner for students.

The embedded zoom tools in the Secure Browser allow students to magnify test content to the following levels (any zoom levels of 5X and greater require users to turn on the Secure Browser's streamlined mode setting, which arranges test content vertically):

- 1X
- 1.5X
- 1.75X
- 2.5X
- 3X
- 5X
- 10X
- 15X
- 20X

Table 5 provides a list of third-party screen magnifier tools that can be used in TDS. The non-embedded screen magnifier tools listed below come with an increased risk of interoperability issues, require students to manually pan the magnification tool across the screen, and can include unwanted features that should not be used while testing. These non-embedded tools also cannot be tracked by Montana OPI when gathering data about students' tool use.

Table 5. Third-Party Screen Magnifier Applications

Product	System Requirements	Additional Details
ZoomText Magnifier • Supported Versions: 2020, 2021, 2022	• Windows 8.1, 10, 11; 2012 R2, 2016 R2	• ZoomText includes a SpeakIt text-to-speech tool that could be used to read aloud passages, which is not permitted on ELA tests. Students testing with ZoomText should use the magnification features only. It is recommended that students requiring text-to-speech (TTS) support use the Secure Browser's embedded TTS tools, and that students requiring screen readers use JAWS or Fusion.

Product	System Requirements	Additional Details
Fusion (combines JAWS screen reader with zoom text) • Supported Versions: 2020, 2021, 2022	• Windows 8.1, 10, 11; 2012 R2, 2016 R2	• Requires additional setup before use with TDS (see configuration instructions for JAWS).

Testing with Text-to-Speech

Text-to-Speech (TTS) tools read aloud text that appears on the screen for students who may have reading impairments. TDS includes embedded TTS tools that can be turned on for students with the appropriate accommodation settings (either in TIDE or from the TA Site). In order for students to test with TTS tools, a supported voice pack will need to be installed on their device before testing begins (if the device does not already include a built-in voice pack). Students testing with TTS should also have a supported headset or headphones.

For more information about enabling English and Spanish voice packs, see the Windows, macOS, and Chrome support pages on this topic.

TTS is available on all operating systems supported by TDS (for a full list of supported operating systems, see the ***Technology Resources*** page of the Montana Testing Portal). However, TTS tracking does not function correctly on Linux devices. If students require the use of this accommodation (TTS with tracking), they must use a different operating system.

Text-to-speech in Windows, macOS, and iPads includes a feature that allows students to pause and then resume TTS in the middle of a passage. On Chromebooks, however, students should highlight the desired text to be read as the pause feature does not allow students to pause and resume reading again.

Table 6 lists the voice packs supported for students testing with TTS. Students can verify that TTS works on their computers by logging in to a practice test session and selecting a test for which TTS is available. Students using TTS for the practice tests must log in using a supported Secure Browser or a supported Chrome, Firefox, or Edge browser.

Table 6. Technology Requirements for Students Testing with TTS

Supported Voice Packs
• Windows ▪ For English tests: All built-in voice packs ▪ For Spanish tests: Spanish (Mexico) • macOS built-in voice packs • iPadOS built-in voice packs • Chromebook built-in voice packs • Heather Infovox iVox HQ (macOS only) • Rosa Infovox iVox HQ (macOS only)

Note: CAI strongly encourages schools to test the TTS settings before students take operational tests. You can check these settings through the diagnostic page. From the student practice test login screen, click **Run Diagnostics**, and then click **Text-to-Speech Check**.

Note: The Mobile Secure Browser uses either the device's native voice pack or a voice pack embedded in the Secure Browser. Additional voice packs downloaded to a mobile device are not recognized by the Mobile Secure Browser.

Testing with Assistive Technology for Braille Tests

Braille tests administered in TDS require the use of multiple assistive technology devices and applications, including the refreshable braille displays (RBDs) and JAWS screen readers used by students to read and navigate test content and the embossers used by proctors to print test content.

RBDs are used to read text-only content on ELA and Mathematics tests, while Braille embossers are needed to read any content with images in ELA tests, as well as advanced content in Mathematics and Science tests. RBDs must be properly setup before they can be used by students. For information about installing and setting up RBDs, refer to the product's provided instructions and manuals.

TDS includes several embedded tools that facilitate braille testing, such as braille presentation settings, various print tools for embossing content, and streamlined mode, which arranges test content vertically.

Note: Alt text for the images in items and stimuli is consistently available only when the test presentation is set to *braille*.

Table 7 provides a list of supported screen reader software that students can use in TDS. **Please note that only JAWS may be used on ELA and Reading tests, as this is the only supported screen reader that can effectively mute reading passages.** Screen readers other than JAWS must not be used on ELA and Reading tests, as they would allow students to listen to passages instead of reading them, compromising the ability to assess their reading comprehension skills.

Table 7. Screen Readers Supported for Student Computers

Screen Reader	System Requirements	Additional Details
JAWS–Student • Supported Versions: 2020, 2021, 2022	• **Operating Systems**: Windows 8.1, 10, 11 ▪ **Minimum Requirements**: 2.0 GHz Processor, 8 GB RAM	• Requires additional setup before use with TDS (see configuration instructions). • Test Presentation setting must be set to braille, whether or not student is a braille user.
VoiceOver	• macOS 13	• Students cannot use VoiceOver for ELA tests, as the read-aloud of passages cannot be suppressed in this product. • Must use Secure Browser 15.5 • Requires additional setup before use with TDS (see configuration instructions).
Fusion • Supported Versions: 2020, 2021, 2022	• **Operating Systems**: Windows 8.1, 10, 11 ▪ **Minimum Requirements**: 2.0 GHz i3 dual core processor, 4 GB RAM	• Requires additional setup before use with TDS (see configuration instructions for JAWS). • Test Presentation setting must be set to braille, whether or not student is a braille user.

Table 8 provides a list of supported refreshable braille displays (RBDs) that students can use to read braille content. Please note that if students wish to use RBDs not mentioned in this table, they should test them on a practice test to ensure there are no issues before using them on an operational test. Additionally, RBDs not listed here may include unwanted features that students should not use while testing, so students may need to be monitored if they use such RBDs.

Table 8. Refreshable Braille Displays Supported for Student Computers

RBD	System Requirements	Additional Details
Brailliant 40 Cell	• Windows 8.1, 10, 11	• CAI recommends RBDs with at least 40 cells, but students may use displays with fewer cells if preferred. • Students should not use the HumanWare Brailliant BI 14 RBD. It can automatically synchronize notes typed internally with a mail application, potentially violating test security.

Table 9 provides a list of embossers and embossing software supported for TA computers. Embossers must be used to print any test content that cannot be read by RBDs, including all content on Mathematics and Science tests, and some of the content on ELA and Social Sciences tests. Different embossing software is required for printing PRN and BRF file types. The printed file types depend on the content being embossed. If your student or school requires embossing software not listed in Table 9, please contact the Assessment Help Desk at the Montana OPI at 1-844-867-2569, as state consultation and approval may be required for additional software.

Table 9. Embossers and Embossing Software Supported for TA Computers

Embosser / Embossing Software	System Requirements	Additional Details
Duxbury Braille Translator • Supported Versions: 12.1, 12.2, 12.5, or 12.6	• **Operating Systems**: Windows 8.1, 10, 11 • **Minimum Requirements**: 1 GHz Processor, 1 GB RAM (for 32-bit), 2 GB RAM (for 64-bit)	• Requires additional setup before use with TDS (see configuration instructions). • Used for embossing BRF files (from print requests containing only text or formatted tables).

Embosser / Embossing Software	System Requirements	Additional Details
ViewPlus Max Embosser, ViewPlus Premier Embosser, ViewPlus Columbia Embosser, ViewPlus Columbia 2 Embosser, or ViewPlus Rogue Embosser PixBlaster	• Windows 8.1, 10, 11	• Requires additional setup before use with TDS (see configuration instructions). • Used for embossing PRN files (from print requests with tactile or spatial components, such as images). • PRN files are formatted for a specific printer driver (e.g., Max embosser). Thus, you may need to convert the PRN file in Tiger Designer for use with another supported embosser (see PRN conversion instructions for more details).
ViewPlus Desktop Embosser (driver for ViewPlus Embossers and Tiger Viewer Software)	• Windows 8.1, 10, 11	• Download and install your embosser driver prior to embossing any files. • The Tiger Viewer software is downloaded in the printer driver folder.
Tiger Software Suite (Tiger Designer) • Supported Versions: 6, 7, 8	• Windows 8.1, 10, 11	• You should download Tiger Designer prior to testing, as some PRN files will need to be converted in this program before embossing. • Please see PRN conversion instructions for more details. • Tiger Software Suite is included with all ViewPlus embossers and its license can be used on up to two devices.

Specifications for TAs Using Screen Readers

If a TA requires the use of a screen reader to set up or administer test sessions in the TA Site, CAI recommends they do so using the most recent Firefox or Chrome browser. If issues occur while updating browsers, please contact your network administrator/IT office.

Configuring JAWS Screen Readers on Student Computers Before Testing Begins

This section includes instructions for the additional JAWS configuration steps that Technology Coordinators must follow before students use JAWS for online testing. To ensure JAWS is properly configured, students should take practice tests using JAWS before taking operational tests.

The configuration instructions in this section apply to JAWS 2020, JAWS 2021, and JAWS 2022 as well as Fusion.

Configuring JAWS to Recognize the Secure Browser

You must edit the JAWS configuration file so that the software recognizes the Secure Browser.

1. To modify the configuration file, open the JAWS **ConfigNames.ini** file. This file may appear in two folders. Depending on how JAWS is installed on your computer, you may need to modify both files. The examples below are for JAWS 2020 installed to the default location. If your version is installed to a different location, navigate to the appropriate directory.

   – **Required Folder**: Start > All Programs > JAWS 2020 > Explore JAWS > Explore Shared Settings

   – **Optional Folder**: Start > All Programs > JAWS 2020 > Explore JAWS > Explore My Settings

2. In the **ConfigNames** file, locate the line of text containing *firefox:3=firefox*. At the end of this line, press **Enter** and type *STATESecureBrowser15.0=firefox*

3. At the bottom of the **ConfigNames** file, do one of the following:

   – **For JAWS 2022**: Update the URL in the *regex:* line below the *;CAI testing:* line so that it reads as *regex:https:\/\/.*\.cambiumtds(uat)?\.com\/student.*=CAI Testing*

   – **For JAWS 2020 and JAWS 2021**: After the last line of the file, press **Enter** and type *;CAI testing:* and then press **Enter** again and type *regex:https:\/\/.*\.cambiumtds(uat)?\.com\/student.*=CAI Testing*

   The end of the **ConfigNames** file should match the highlighted section in Figure 1.

   Figure 1. JAWS ConfigNames File—regex line

4. Save the file.

   a. If you receive an error that you don't have permission to save the .ini file to this location, save the file as `ConfigNames.ini` to your desktop. Then copy the updated .ini file to the folder containing the original .ini file referenced in step 1.

Additional Configuration for JAWS 2020

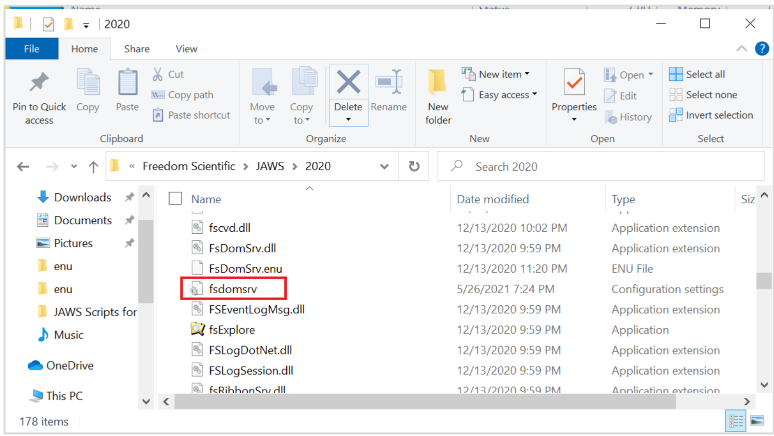
There is a known issue in which older versions of JAWS do not run properly with the Secure Browser. To ensure these versions of JAWS function properly, you must download an additional JAWS file and replace the existing version of that file that appears in your JAWS settings folder. These steps do not need to be performed for JAWS 2021 or later.

1. Click here to download the **fsdomsrv.ini** zip file and save it to your desktop.

2. Locate the JAWS folder that contains the existing **fsdomsrv** file:

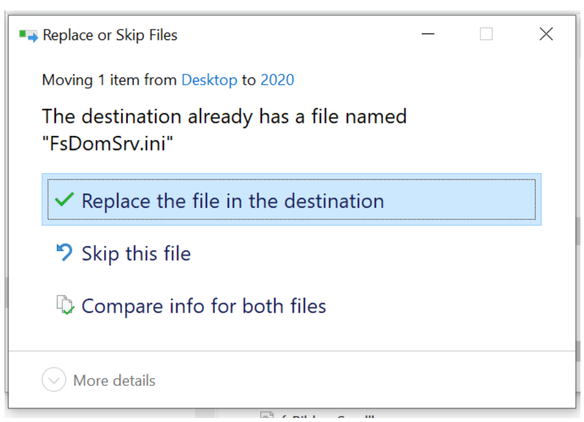
   – **Folder location for JAWS 2020**: Local Disk > Program Files > Freedom Scientific > JAWS > 2020

Figure 2. Original fsdomserv File Location



3. Drag and drop the newly downloaded **fsdomsrv.ini** file from the desktop into the folder listed in step 1. In the ***Replace or Skip Files*** window, select **Replace the file in this destination**.

Figure 3. Replace or Skip Files Window



Applying Settings for Contracted or Uncontracted Braille

In order for students to use contracted or uncontracted literary braille with their RBD, the correct JAWS settings must be applied prior to launching the Secure Browser.

1. To apply the correct JAWS settings, open JAWS and go to **Utilities** > **Settings Center**. The ***Settings Center*** window opens.

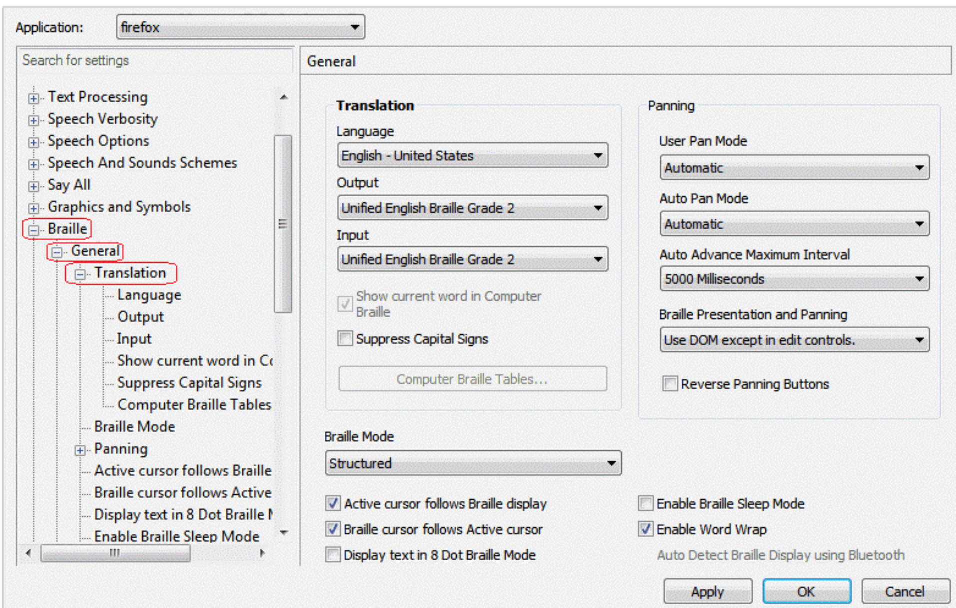
2. From the **Application** drop-down list at the top of the window, select **firefox**.

3. Expand the *Braille* settings, *General* sub-settings, and *Translation* sub-settings in the *Search for settings* panel on the left. The ***Settings Center*** window displays the options for Braille Translation (see [Figure 4](#)).

   a. In the *Translation* section, verify the **Language** drop-down list is set to **English – United States**. For a student who prefers contracted braille, select **Unified English Braille Grade 2** from the **Output** and **Input** drop-down lists. For a student who prefers uncontracted braille, select **Unified English Braille Grade 1** from the **Output** drop-down list.

      i. For tests presented in the EBAE braille type, if the student prefers contracted braille, select **US English Grade 2** from the **Output** and **Input** drop-down lists. If the student prefers uncontracted braille, select **US English Braille Grade 1** from these drop-downdrop-down lists.

Figure 4. JAWS Settings Center Window



4. In the *Braille Mode* section (see [Figure 5](#)), ensure that only the following settings are checked:

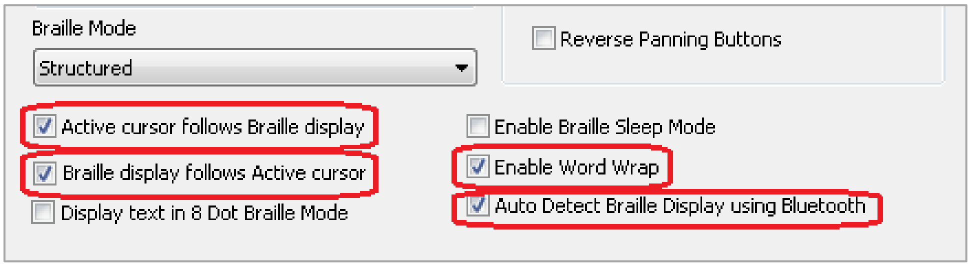
   – Active cursor follows Braille display

   – Braille cursor/display follows Active cursor

   – Enable Word Wrap

   – Auto Detect Braille Display using Bluetooth (if available)

Figure 5. Braille Mode Section



5. Click **Apply**, and then click **OK**.

Note: When students testing with contracted braille type a written response with their RBD, the words will be parsed while typing, so each word may not display on screen until after pressing **Space** or **Enter**. This is the intended behavior when using this setting.

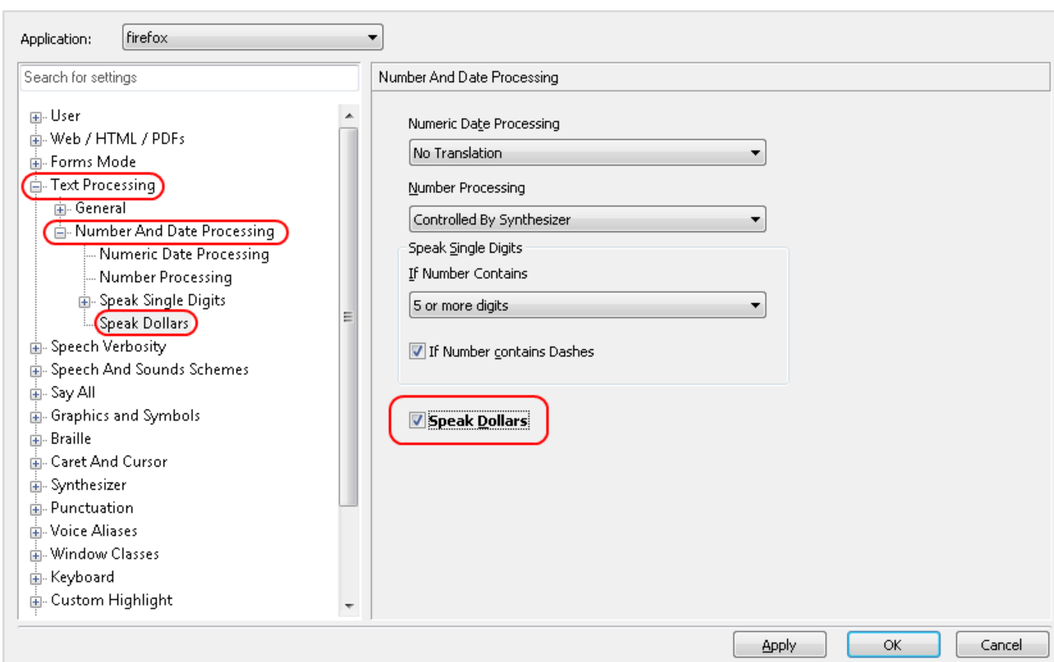
Configuring JAWS to Speak "Dollars"

If a test includes content with the dollar symbol ($), you should configure JAWS to correctly speak this symbol.

1. Open JAWS and go to **Utilities** > **Settings Center**. The ***Settings Center*** window opens.

2. In the *Search for settings* panel on the left, expand the *Text Processing* settings and *Number And Date Processing* sub-settings. Click **Speak Dollars**. The ***Settings Center*** window displays the *Number And Date Processing* options (see Figure 6).

Figure 6. Number and Date Processing



3. Mark the **Speak Dollars** checkbox.

4. Click **Apply**, and then click **OK**.

JAWS Unified Keyboard Settings

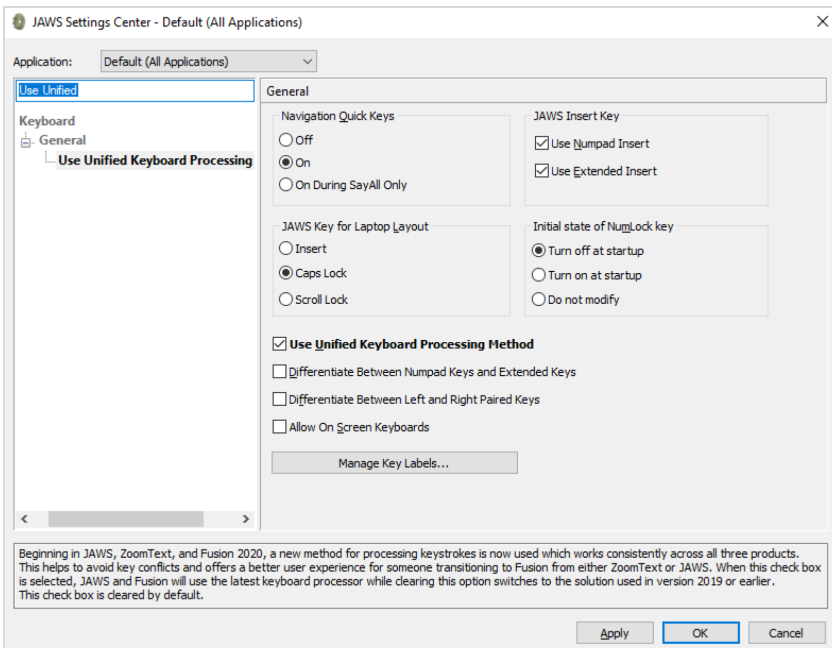
JAWS includes a unified keyboard setting that allows students to use **Alt**+**Tab** to return to the Secure Browser when Permissive Mode is turned on.

The unified keyboard setting is turned on by default, but you should still verify it is enabled for students before they begin testing.

1. Open JAWS and navigate to **Utilities > Setting Center**.

2. Search in the ***Settings Center*** window for "Unified Keyboard" and mark the **Use Unified Keyboard Processing Method** checkbox (see Figure 7).

Figure 7. Use Unified Keyboard Processing Setting

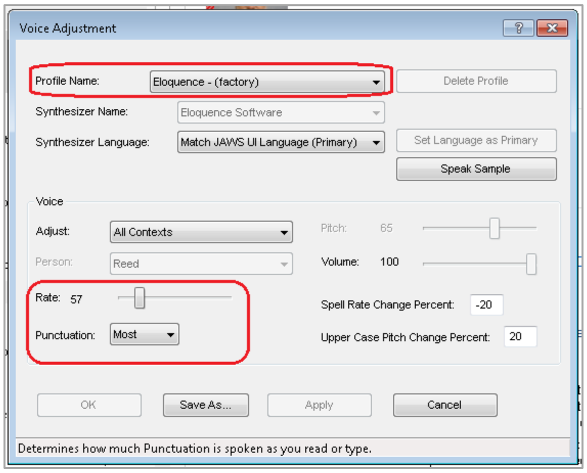


Optional JAWS Voice Adjustment Settings

Prior to launching the Secure Browser, you can adjust JAWS voice settings for students based on their individual needs. You must set the Voice Profile, Speaking Rate, and Punctuation settings prior to administering assessments. Students should take practice tests using JAWS so they can determine whether these settings need to be adjusted.

1. To adjust JAWS voice settings, open JAWS and go to **Options** > **Voices** > **Voice Adjustment**. The ***Voice Adjustment*** window opens (see Figure 8).

Figure 8. JAWS Voice Adjustment



2. To adjust the voice profile, in the *Profile* section, select a voice profile from the **Profile Name** drop-down list. Click **Apply**.

3. To adjust the voice rate, in the *Voice* section, drag the **Rate** slider to the desired rate speed (the lower the rate, the slower the words are read aloud). Click **Apply**.

4. To adjust the punctuation, click the **Punctuation** drop-down list. Select from the following options: **None**, **Some**, **Most**, or **All**. Click **Apply**.

5. When all settings are saved, click **OK**.

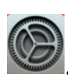
Configuring VoiceOver Screen Readers on Student Computers Before Testing Begins

This section includes instructions for the additional VoiceOver configuration steps that Technology Coordinators must follow before students use VoiceOver for online testing. To ensure VoiceOver is properly configured, students should take practice tests using VoiceOver before taking operational tests.

Students cannot use VoiceOver for ELA tests, as it cannot suppress the read-aloud of passages.

Launching VoiceOver Through Accessibility Settings

Prior to testing with VoiceOver, you can configure its settings by launching the software through the Accessibility Settings center on macOS devices.

1. In the dock, click **System Settings** .

2. Select **Accessibility** > **Shortcut** (in the *General* section).

3. Mark the **VoiceOver** checkbox.

4. You can now quickly enable VoiceOver from the Accessibility Shortcuts window by pressing **Command** + **Option** + **F5** and then selecting **VoiceOver**.

Using the VO Key for VoiceOver Commands

Throughout this section, many VoiceOver instructions mention a "**VO**" key. Depending on how VoiceOver is set up on the student's computer, the **VO** key may either be the **Control** + **Option** keys pressed simultaneously or the **Caps Lock** key. For example, if a section in this guide instructs you to press **VO** + **F8**, you would press either **Control** + **Option** + **F8** or **Caps Lock** + **F8**.

You can specify which command should act as the **VO** key from the ***VoiceOver Utility*** window:

1. To turn on VoiceOver, press **Command (⌘) + F5**. Then press **VO** + **F8** to open the ***VoiceOver Utility*** window.
2. In the left panel, select **General**.
3. From the VoiceOver modifier drop-down list, select the keys you want to use as the **VO** key:
   - **Control** + **Option** (^⌥)
   - **Caps Lock**
   - Either **Control** + **Option** or **Caps Lock**

Note: If you set **Caps Lock** as the **VO** key, students who wish to activate Caps Lock and type uppercase characters should quickly press **Caps Lock** twice.

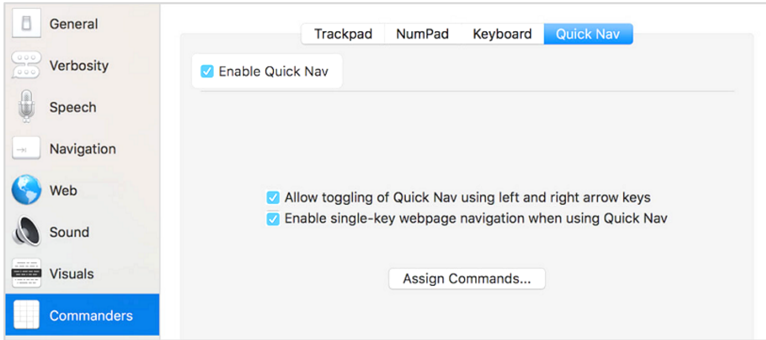
Configuring VoiceOver Screen Reader Navigation Settings

VoiceOver allows students to use keyboard commands to navigate the Student Testing Site. For a better experience when using keyboard commands, you should enable the ***Quick Nav*** feature and other settings in the ***VoiceOver Utility*** window.

Note: CAI does not recommend the use of numpad commands for navigating tests due to compatibility issues. For the best experience, students should use the keyboard navigation mode.

1. To turn on VoiceOver, press **Command** + **F5**. Then press **VO** + **F8** to open the ***VoiceOver Utility*** window.
2. In the left panel, select **Commanders**.
3. On the **Quick Nav** tab, make sure the following checkboxes are marked:
   - **Enable Quick Nav**
   - **Allowing toggling of Quick Nav using left and right arrow keys**
   - **Enable single-key webpage navigation when using Quick Nav**

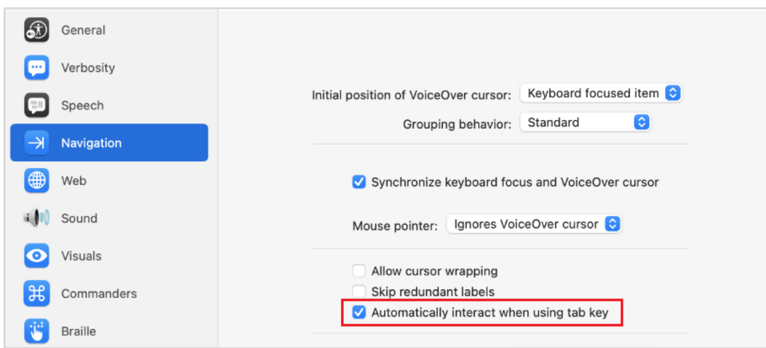
Figure 9. Commanders – Quick Nav Tab



4. To ensure that the tab key functions properly when using VoiceOver in a test, select **Navigation** in the left panel. Then make sure the following checkbox is marked:

   – **Automatically interact when using tab key**

Figure 10. Navigation Panel



5. To ensure that VoiceOver interacts properly with websites, select **Web** in the left panel and then click the **Web Rotor** tab. Make sure the following checkboxes are marked:

   – **Links**

   – **Headings**

   – **Form Controls**

   – **Tables**

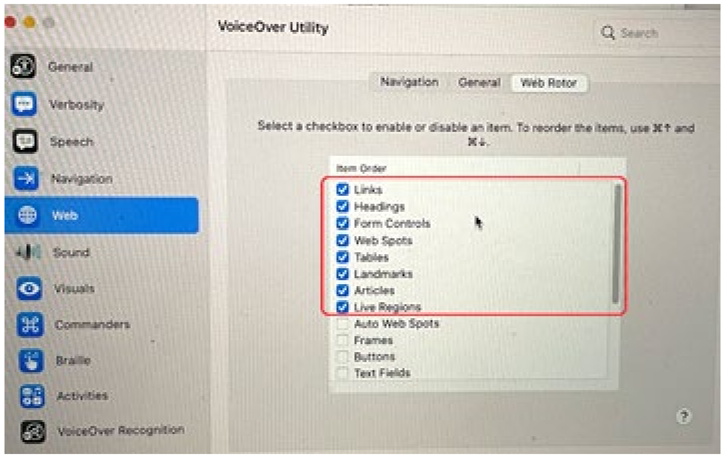
   – **Landmarks**

   – **Frames**

   – **Web Spots**

   – **Live Regions**

Figure 11. Web – Web Rotor Tab



Applying VoiceOver Settings for Contracted or Uncontracted Braille

In order for students to use contracted or uncontracted literary braille with their RBD, you must apply the correct VoiceOver settings prior to launching the Secure Browser.

1. To turn on VoiceOver, press **Command** + **F5**. Then press **VO** + **F8** to open the ***VoiceOver Utility*** window.

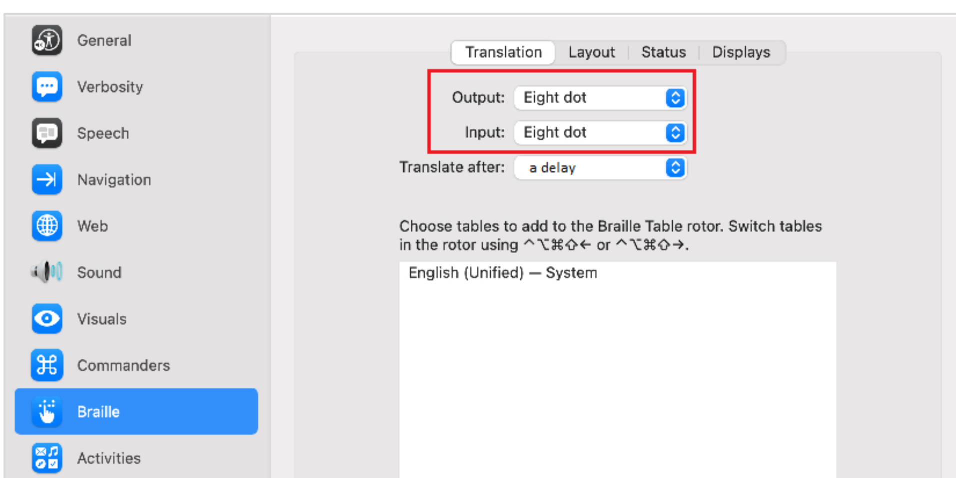
2. In the left panel, select **Braille**.

3. In the *Translation* tab, do one of the following:

   – For a student who prefers contracted braille, select **Contracted six dot** from the **Output** and **Input** drop-down lists.

   – For a student who prefers uncontracted braille, select either **Six dot** or **Eight dot** from the **Output** and **Input** drop-down lists.

Figure 12. Braille – Translation Tab



4. On the **Layout** tab, ensure that the following settings are checked:

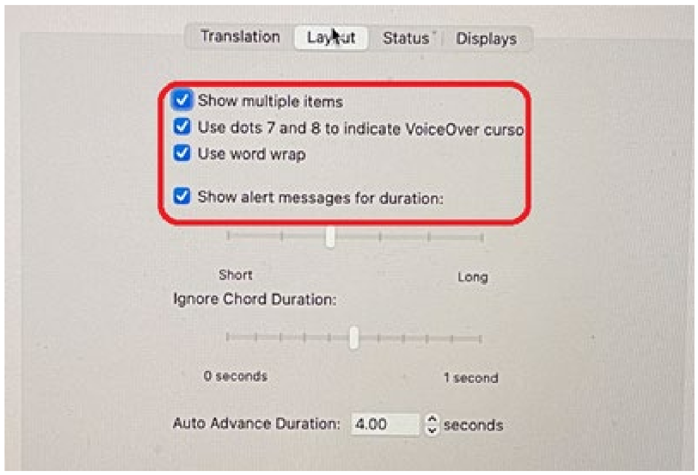
   – **Show multiple items**

   – **Use dots 7 and 9 to indicate VoiceOver Cursor**

   – **Use word wrap**

   – **Show alert messages for duration**

Figure 13. Braille – Layout Tab



Optional VoiceOver Voice Adjustment Settings

Prior to launching the Secure Browser, you can adjust VoiceOver voice settings based on each student's individual needs. You must set the Voice Profile and Speaking Rate settings prior to administering assessments. Students should take practice tests using VoiceOver so they can determine whether these settings need to be adjusted.

1. To turn on VoiceOver, press **Command** + **F5**. Then press **VO** + **F8** to open the ***VoiceOver Utility*** window.

2. In the left panel, select **Speech**.

3. On the **Voices** tab, select a voice profile from the **Voice** drop-down list.

4. To adjust the voice rate, set the *Rate* field to the desired rate speed (the lower the rate, the slower the words are read aloud).

Configuring Embossing Software on TA Computers Before Testing Begins

TDS allows students to emboss test material with Test Administrator (TA) approval. The software that sends print requests to the braille embosser must be installed on computers that TAs use for test sessions.

The embossed output for student print requests depends on the file type associated with a test question. TAs must ensure that students have the Braille Type test setting prior to approving the student for testing, as this determines which file type is used for printing. There are two types of files:

- **Braille Ready File (BRF)**: BRF file types are used for print requests containing only text (including formatted tables). The Duxbury Braille Translator software handles BRF files.
- **Printer Output File (PRN)**: PRN file types are used for print requests containing tactile or spatial components (such as images). The ViewPlus software handles PRN files.

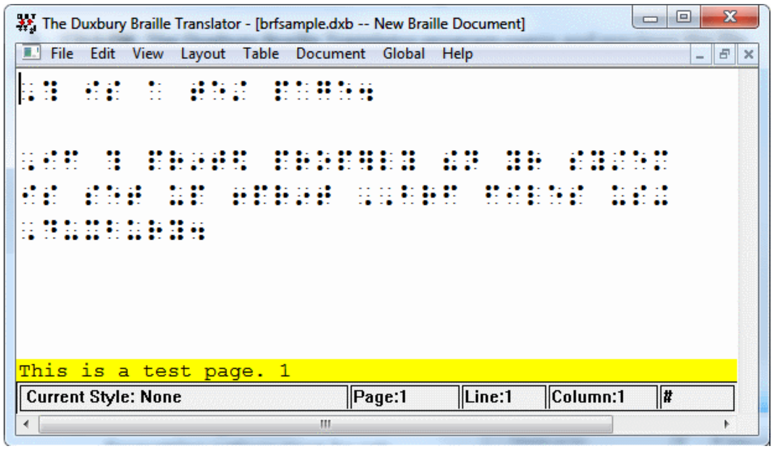
Upon approving a print request, the TA sends the file to the embosser using either Duxbury or ViewPlus software. Instructions for embossing files are located in the section Embossing Braille Print Requests for BRF and PRN Files.

Configuring Duxbury Braille Translator for BRF Files

This section contains instructions for opening BRF files with Duxbury Braille Translator (DBT) and setting default embossing preferences. The DBT software must be installed before performing these steps.

1. In the TA Site, click **Help Guide** at the top of the page. The online *TA User Guide* opens.
   a. Sample braille files can be accessed from the **Help Guide** > **Appendices** > **Sample Braille Files**.
2. Click **Sample BRF File**. The browser downloads the file.
3. Click the downloaded file and do one of the following:
   – From the **More apps** list, select **Duxbury Braille Translator**. Click **OK**. The Duxbury Braille Translator program opens and previews the file (see Figure 14).
   – If the Duxbury Braille Translator is not available as a selectable program, do the following (otherwise skip to step 4):
      i. Click **Look for another app on this PC**. The ***Choose Helper Application*** window opens.
      ii. Navigate to the Duxbury folder (in the Program Files (x86) folder) and open it.
      iii. Open the DBT folder and select **dbtw.exe**.
      iv. Click **OK**. The Duxbury Braille Translator program opens and previews the file (see Figure 14).
      v. If the ***Import File*** window appears, set the Template to `English (UEB) – basic`, and set the Import Filter to `Formatted Braille`.

Figure 14. Duxbury Braille Translator Window



4. In the ***Duxbury Braille Translator*** window, go to **Global** > **Embosser Setup**. The ***Global: Embosser Setup*** window appears. To add a new embosser, do the following:

    a. Click **New**. The ***Embosser Setup – Untitled Configuration*** window appears.

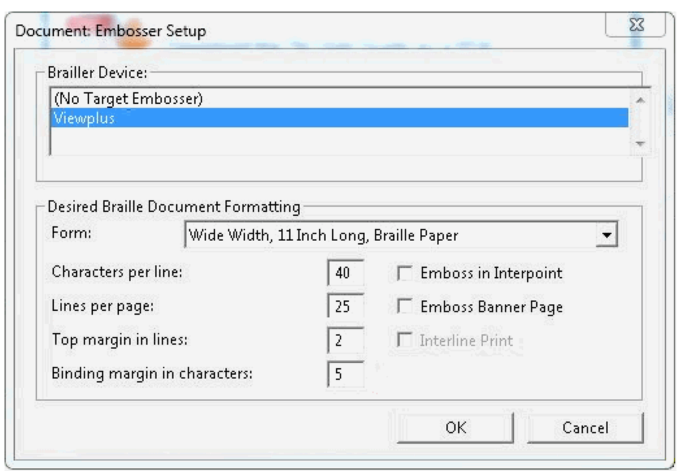
    b. From the **Embosser Model** drop-down list, select the required embosser type.

    c. From the **Send to Printer** drop-down list, select the required embosser's name and click **OK**.

    d. In the ***Global: Embosser Setup*** window, click **OK**.

5. In the ***Duxbury Braille Translator*** window, go to **Document** > **Embosser Setup**. The ***Document: Embosser Setup*** window opens (see Figure 15).

Figure 15. Document: Embosser Setup Window



6. In the ***Document: Embosser Setup*** window, ensure the following are selected:

    – *Brailler Device*: **ViewPlus Max** (or whichever supported ViewPlus embosser you are using)

    – The following *Braille Document Formatting* options must be set:

  - **Emboss in Interpoint** checkbox is blank
  - *Top margin in lines*: 2
  - *Binding margin in characters*: 5
- When you are done, click **OK**.

7. In the ***Duxbury Braille Translator*** window, go to **Global** > **Formatted Braille Importer**.
   a. In the ***Global: Formatted Braille Importer*** window that appears, mark the **Read formatted Braille without interpretation** checkbox and click **OK**.
8. In the ***Duxbury Braille Translator*** window, go to **File** > **Emboss**. The ***File: Emboss…*** window opens.
9. In the ***File: Emboss…*** window, ensure that only one copy is being printed and that the page range is set to **All**.
10. Click **OK**.

Configuring ViewPlus Tiger Designer or Tiger Viewer for PRN Files

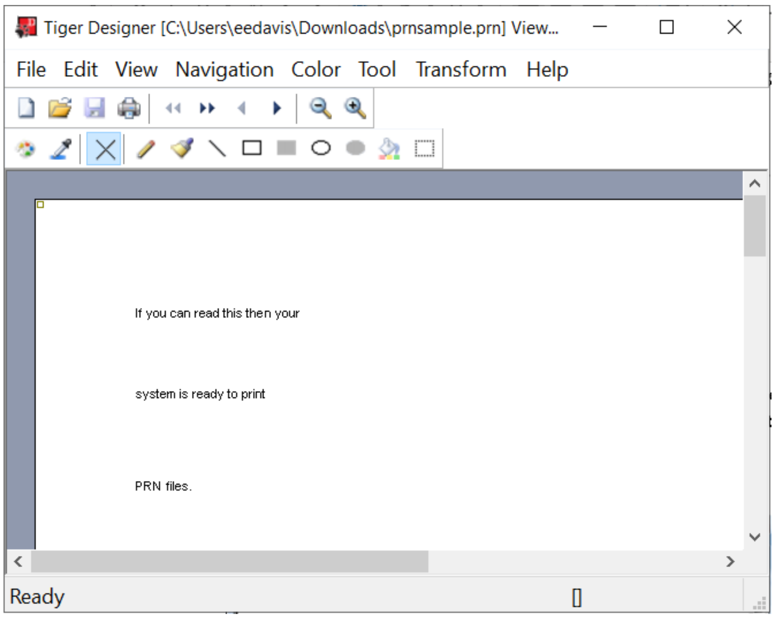
This section contains instructions for opening PRN files with ViewPlus software and setting default application preferences.

Although you may use Tiger Viewer as the default application for opening PRN files, if you use it alongside an embosser other than the Viewplus Max or Premier, you may need to convert the files for your embosser type before printing. Because of this, CAI strongly recommends using Tiger Designer as the default application for embossing PRN files.

1. In the TA Site, click **Help Guide** at the top of the page. The online *TA User Guide* opens.
   a. Sample braille files can be accessed from the **Help Guide** → **Appendices** → **Sample Braille Files**.
2. Click **Sample PRN File**. The browser downloads the file.
3. Click the downloaded file and do one of the following:
   - To set Tiger Designer as the default application (recommended):
     - From the **More apps** list, select **Tiger Designer** and click **OK**. The Tiger Designer program opens and previews the file (see Figure 16).
     - If Tiger Designer is not available as a selectable program, click **Look for another app on this PC** and select Tiger Designer from the folder where it is installed on your computer.
   - To set Tiger Viewer as the default application:
     - From the **More apps** list, select **Tiger Viewer** and click **OK**. The Tiger Viewer program opens and previews the file.

- If Tiger Viewer is not available as a selectable program, click **Look for another app on this PC** and select Tiger Viewer from the folder where it is installed on your computer.

Figure 16. Tiger Designer Window



6. Go to **File** > **Print**. The ***Print*** window opens.

7. Ensure that the printer is set to **ViewPlus Max** (or whichever supported ViewPlus embosser you are using) and that only one copy is being printed.

8. Click **Print**.

   – If you experience any issues embossing, refer to the section [Converting PRN Files in Tiger Designer for Embossing]() or contact the Help Desk.

Administering Braille Tests

This section explains how TAs set up the test settings for braille tests and emboss braille print requests from students. It also provides information about how students navigate the Secure Browser with JAWS.

Setting Up Braille Test Sessions

TAs must make sure that students have the correct test settings applied before approving them to take braille tests. Any test settings that cannot be changed from the TA Site or Secure Browser will need to be set in TIDE. Please note that some test settings may vary between practice and operational tests.

For more detailed instructions about starting test sessions, see the *Test Administrator User Guide*.

1. To administer braille tests, the TA logs in to the appropriate TA Site and starts a test session.

2. The TA opens the screen readers on the student testing devices.

3. The TA opens the Secure Browser on the student testing devices.

4. Students sign in to the test session and select their tests.

5. The TA reviews the student's test settings and verifies the following:

   – *Presentation* is set to **Braille**. This should be set for any students testing with a screen reader, regardless of whether those students are braille users. Setting the Presentation to **Braille** will automatically enable streamlined mode, which arranges test content vertically.

   – *Print on Request* is set to the appropriate option for the selected test.

   – *Braille Type* is set to the student's preferred braille option. Students may choose from the following options, depending on the test:

English Language Arts	Mathematics	Science
▪ UEB Contracted with Nemeth Math ▪ UEB Uncontracted with Nemeth Math	▪ UEB Contracted with Nemeth Math ▪ UEB Uncontracted with Nemeth Math ▪ UEB Contracted with UEB Math ▪ UEB Uncontracted with UEB Math	▪ UEB Contracted with Nemeth Math

   – *Emboss Request Type* is set to **Auto-Request** or **On-Request**, depending on the rules for the selected test.

   – *Mute System Volume* is set to the appropriate option for the student and the screen reader that the student is using. This setting prevents JAWS from reading aloud passages on ELA tests.

     ▪ This setting is available only for JAWS screen readers, which is why JAWS is the only screen reading software supported for ELA tests with reading passages.

   – *Audio Transcriptions* is set to the appropriate option for the student. When this tool is enabled, any audio content in the test will have an associated transcript in the global menu that can be read by the student's RBD.

   – *Permissive Mode* is turned on. This setting must be enabled in order for students to use the keyboard commands associated with their screen reader.

6. When all the correct settings are applied, the TA approves students for testing.

Embossing Braille Print Requests for BRF and PRN Files

This section provides instructions for TAs who will emboss BRF and PRN files during a test session. As students progress through their tests, emboss requests are sent to the TA Site, either automatically or manually, depending on the test settings. You must review and approve these emboss requests in order to send the files to the embossers. BRF files may be embossed with Duxbury Braille Translator. PRN files may be embossed with either Tiger Designer or Tiger Viewer. However, Tiger Designer is the recommended embosser, as it also allows you to convert file types if necessary.

When the test session is over, you must delete and discard all test materials. This may require you to remove files from the web browser download archive.

For tips and troubleshooting advice for embossing braille files, see Appendix A.

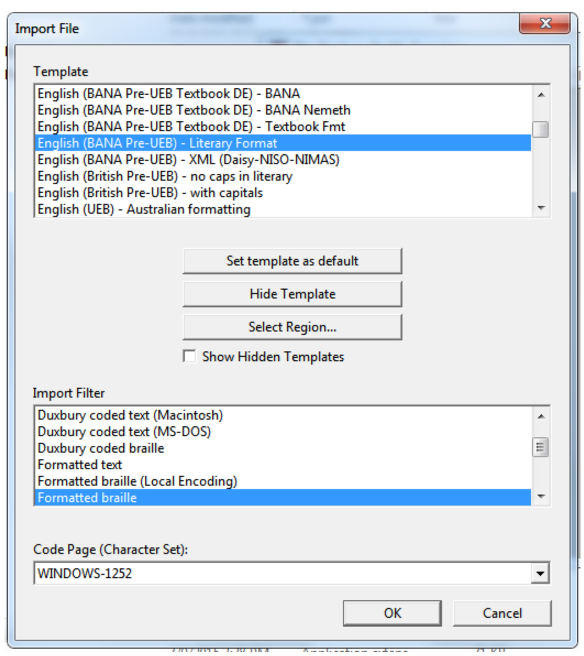
Sending BRF Files to the Embosser

BRF files must be embossed with Duxbury Braille Translator. For information about setting up Duxbury Braille Translator prior to testing, see the section Configuring Duxbury Braille Translator for BRF Files.

1. When you approve a print request for a BRF file, either open the file directly from the browser or save it and open it in Duxbury Braille Translator. See Appendix A for more information about saving and opening print requests.

    a. If you opened the file directly from your browser and Duxbury Braille Translator is the default application for printing BRF files, the ***Import File*** window appears (see Figure 17). If not, launch Duxbury Braille Translator and then select **File** > **Open** to select the downloaded BRF file.

Figure 17. Import File Window



2. Ensure that the following settings are selected:

   – Template: **English (UEB) – basic**

   – Import Filter: **Formatted braille**

3. Click **OK**. The ***Duxbury Braille Translator*** preview window opens (see Figure 14).

4. Go to **File** > **Emboss**. The ***File: Emboss*** window opens.

5. Ensure that only one copy is being printed, the page range is set to **All**, and the *Brailler Device* is set to **ViewPlus Max** (or other ViewPlus embosser). Then click **OK**.

Sending PRN Files to the Embosser with Tiger Designer

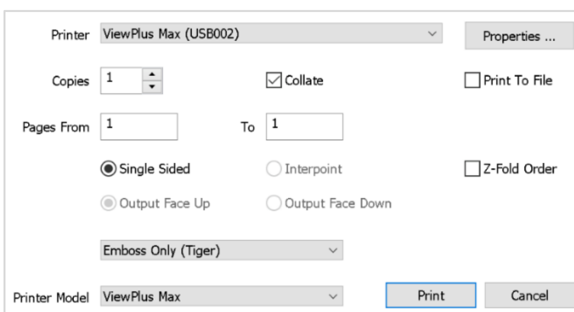
Tiger Designer is the recommended software for embossing PRN files. For information about setting up Tiger Designer prior to testing, see the section Configuring ViewPlus Tiger Designer or Tiger Viewer for PRN Files.

1. When you approve a print request for a PRN file, either open the file directly from the browser or save it and open it in Tiger Designer. See Appendix A for more information about saving and opening print requests.

   a. If you opened the file directly from your browser and Tiger Designer is set as the default program for PRN files, a ***Print*** window should appear automatically. If not, launch Tiger Designer and select **File > Print** to select the downloaded PRN file.

2. Ensure that only one copy is being printed and the *Printer Name* is set to **ViewPlus Max** (or whichever supported ViewPlus embosser you are using), then click **Print**.

Figure 18. Tiger Designer Print Window



Sending PRN Files to the Embosser with Tiger Viewer

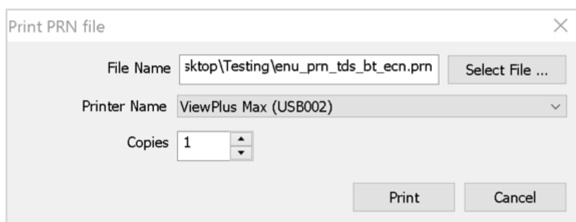
You can also emboss PRN file print requests in Tiger Viewer. However, if the print option is grayed out when printing a file, you will not be able to convert the file with Tiger Viewer. PRN files can only be converted in Tiger Designer. For information about setting up Tiger Viewer prior to testing, see the section Configuring ViewPlus Tiger Designer or Tiger Viewer for PRN Files.

1. When you approve a print request for a PRN file, either open the file directly from the browser or save it and open it in Tiger Viewer. See Appendix A for more information about saving and opening print requests.

    a. If you opened the file directly from your browser and Tiger Viewer is the default application for printing PRN files, the ***Print PRN File*** window appears. If not, launch Tiger Viewer and then select **File** > **Print File** to select the downloaded PRN file.

2. Ensure that only one copy is being printed and that the *Printer Name* field is set to the embosser that you will be using to emboss the braille files. Then click **Print**.

Figure 19. Print PRN File Window



- If the option to **Print** is grayed out, you will need to convert the file in Tiger Designer.

Converting PRN Files in Tiger Designer for Embossing

If you are using an embosser other than VP Max or Premier, you may encounter PRN embossing issues that cause the Tiger Viewer ***Print PRN File*** window or the Tiger Designer ***Print*** window to appear grayed out. In most cases, you can follow the steps below to resolve this issue. If these steps do not work, please consult the Help Desk.

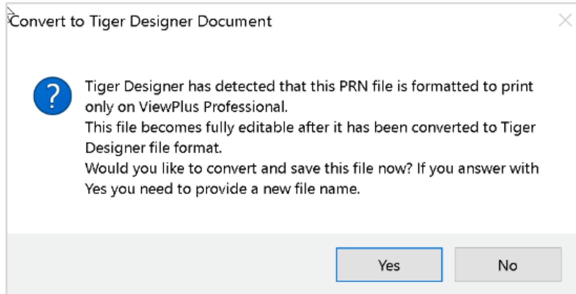
1. Launch Tiger Designer.

2. Go to **File** > **Open**. Then select the file that is not embossing properly and click **Open**.

    a. You may need to use the drop-down above the **Open** button to change the file type to **Tiger PRN files (*.prn)** before you open it.

3. Some improperly formatted PRN types will trigger the message shown in Figure 20.

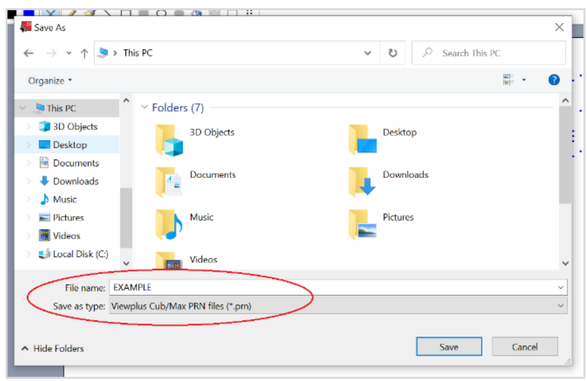
Figure 20. Convert File Message



– If this popup message appears, click **Yes**.

– If this popup message does not appear, navigate to **File > Save As**. Choose the destination where you wish to save the file, and then select **Viewplus Cub/Max PRN files (*.prn)** from the **Save as Type** drop-down (see Figure 21).

Figure 21. Resaving the PRN File



4. Exit Tiger Designer, then reopen the newly saved PRN file in Tiger Designer.

5. In the top toolbar, select **File > Print**.

   a. A ***Print*** window appears. Ensure that only one copy is being printed and the *Printer Name* is set to **ViewPlus Max** (or whichever supported ViewPlus embosser you are using), then click **Print**.

Removing Files from the Web Browser Download Archive

Most supported web browsers automatically save downloaded files. If your computer saves the BRF and PRN files from print requests, for security purposes, you must delete all test-related files from your browser's download archive.

To remove files in Google Chrome:

1. Open the Chrome menu icon in the top-right corner.

2. Select **Downloads**. The ***Downloads*** page opens.

3. Remove all test-related files by doing one of the following:

- For each file, click **X**.
- Click **Clear all** in the top-right corner. Files saved to your computer are not deleted.

To remove files in Edge:

1. Click the Edge Hub (favorites, reading list, bookmarks and downloads) icon in the upper-right corner.
2. Select **Downloads** from within the downloads list.
3. Select each file and click **X** to delete it.

To remove files in Mozilla Firefox:

1. Open the **Tools** menu and select **Downloads**. The ***Library*** window opens.
2. Delete all test-related files by doing one of the following:
   - Select each file and press **Delete** on your keyboard.
   - Click **Clear Downloads** at the top of the window (if available). Files saved to your computer are not deleted.

Navigating the Student Testing Site with JAWS

JAWS allows students to use keyboard commands to navigate the Student Testing Site. Students using RBDs with router keys may also press the router key above the text for a button to move the cursor to that button. They can press the router key again to select that button instead of using the provided keyboard commands.

The actions associated with each JAWS keyboard command depend on the context in which the student presses the key. In other words, the same key may have different effects depending on whether the student is on the sign-in pages, the test pages, or within the items and stimuli of the test pages.

Table 10 provides an overview of how to use JAWS keyboard commands in each context. In order for students to use these keyboard commands, Permissive Mode must be enabled for them in TDS. If JAWS enters Forms Mode, these keyboard commands may not work. In order to exit Forms Mode, press **NUM PAD PLUS** or **Esc**.

Table 10. Overview of JAWS Keyboard Commands in the Student Testing Site

Key	Action
	Navigating the Sign-In Pages with JAWS Keyboard Commands
JAWS modifier key + F10) **Space + S** (Perkins Braille keyboard)	Returns the focus to the Secure Browser if the student navigates to the JAWS application window while signing in Keyboard layouts may vary by device. Please refer to the manual provided by the device manufacturer for more information.
Tab	Moves the focus to the next field or button on the page
Shift + Tab	Moves the focus to the previous field or button on the page
Down Arrow	Reads the next line on the page
Up Arrow	Reads the previous line on the page
Enter	Selects the button that is currently in focus
	Navigating Test Pages with JAWS Keyboard Commands

R	Navigates to the next landmark region on the test page. A test page has up to three primary landmark regions:
	• **Banner Region**: The banner contains the test information row. This row displays the current question numbers, test name, student name, test settings button, pause button, and help button.
	• **Navigation and Test Tools Region**: This region displays the navigation and tool buttons.
	• **Test Content Region**: This region consists of the *Stimulus* section and the *Question* section:
	  ▪ *Stimulus Section*: Contains the stimulus title, stimulus context menu, and stimulus content.
	  ▪ *Question Section*: Contains a question number, question labels (labels that appear when you mark an item for review, print an item, or enter a note for an item), question context menu, question prompt, and the response area.

Key	Action
H	Jumps to the next heading on the page. In general, the following test components are defined with a heading: • Test name (H1) • Student name (H2) • Passage title (H3) • Question number (H3) On test pages that have multiple questions, students can jump directly from one question to the next. To do so, press **H** and then press the **Down arrow** twice. The question prompt is read aloud.
Shift + R	Jumps to the previous region on the page.
Shift + H	Jumps to the previous heading on the page.
Tab	Moves to the next component on the page. In general, the following test elements are components: • Navigation and tool buttons • Question number (and associated prompt text) • Context menu • Response options
Shift + Tab	Moves to the previous component on the page.
Enter	Selects a button or response option or opens a context menu.
Down Arrow	Moves to the next line on the page.
Up Arrow	Moves to the previous line on the page.
JAWS modifier key + Down Arrow	Reads everything on the page (from the current point of focus).
Ctrl	Stops JAWS from reading.
Opening and Using Context Menus with JAWS Keyboard Commands	
Enter	Pressing **Enter** when JAWS reads “Menu button” will open the context menu. This is the only way to open the context menu when streamlined mode is turned on.
Down Arrow	Moves the focus to the next option in the menu. JAWS will read this option aloud.

Key	Action
Up Arrow	Moves the focus to the previous option in the menu. JAWS will read this option aloud.
Space	Selects the menu option currently in focus.
Esc	Closes the context menu without selecting any options.
Responding to Items with JAWS Keyboard Commands	
Tab	• Students can use the **Tab** key to navigate to the item prompt, which JAWS will read aloud. • After JAWS reads the prompt aloud, students can press **Tab** again to navigate to the response area. They may need to press **Tab** multiple times depending on the item type and whether any question labels appear for the item. • In the response area for an item, students can press **Tab** to navigate between each answer option, text box, selectable text field, keypad button, or check box, depending on the item type.
Shift + Tab	Navigates to the previous answer option, text box, selectable text field, keypad button, or check box, depending on the item type.
Up and **Down** Arrow Keys	• For multiple choice and multi-select items, pressing the arrow keys will move between each answer option. • For edit task choice items, pressing the arrow keys will move between each line of text in the item. After users open an edit menu by pressing **Space**, the arrow keys can be used to move between the answer options in the drop-down list.
Space	• For multiple choice and multi-select items, pressing **Space** will select the answer option in focus. • For edit task items, pressing **Space** will open the edit menu in which students type or select a response. • For table match items, pressing **Space** will mark the checkbox in focus.
Enter	• For hot text items, pressing **Enter** will choose the selectable text area in focus as the answer option. • For edit task choice items, pressing **Enter** will select an answer option from the drop-down list in the edit menu. • For equation items, pressing **Enter** will select the keypad button in focus.
Alt + **7**	• For equation items, pressing **Alt** + **7** will open a popup menu with special characters. Students can use the arrow keys to move between the special characters in the list and then press **Enter** to insert a special character in the response area.

Please note that the keypad of the Desmos calculator tool in the Student Testing Site cannot be navigated entirely with the **Tab** key. However, all buttons within the keypad can be navigated or selected using JAWS or Desmos keyboard commands. For a list of Desmos keyboard commands, see https://www.desmos.com/graphingshortcuts.

Navigating the Student Testing Site with VoiceOver

VoiceOver allows students to navigate tests using keyboard commands. Some of these commands refer to the **VO** key. For more information about the **VO** key, see Using the VO Key for VoiceOver Commands.

Key	Action
Navigating the Sign-In Pages with VoiceOver Keyboard Commands	
• **VO** + **F1** (x2) (Standard keyboard) • **Space** + **S** (Perkins Braille keyboard)	Returns the focus to the Secure Browser if the student navigates to the VoiceOver application window while signing in. Keyboard layouts may vary by device. Please refer to the manual provided by the device manufacturer for more information.
Tab	Moves the focus to the next field or button on the page.
Shift + **Tab**	Moves the focus to the previous field or button on the page.
VO + **Shift** + **Down Arrow**	Moves focus to interact with web content/elements.
VO + **Shift** + **Up Arrow**	Stops interacting with web content/elements.
VO + **Right Arrow**	Reads the next line/item on the page.
VO + **Left Arrow**	Reads the previous line/item on the page.
Navigating Test Pages with VoiceOver Quick Nav Keyboard Commands	
C	For hot text items, navigates to the next checkbox on the page.
H	Jumps to the next heading on the page. In general, the following test components are defined with a heading: • Test name (H1) • Student name (H2) • Passage title (H3) • Question number (H3) On test pages that have multiple questions, students can jump directly from one question to the next. To do so, press H and then press VO + Right Arrow. The question prompt is read aloud.
C or **Shift** + **C**	For hot text items, navigates to previous checkbox on the page.
H or **Shift** + **H**	Jumps to the previous heading on the page.
B	Moves to next button on the page.
B or **Shift + B**	Moves to previous button on the page.
J	For multiple choice items, navigates to next answer choice.

Key	Action
J or **Shift** + **J**	For multiple choice items, navigates to the previous answer choice.
F	For equation items and text response items, navigates to the next text entry field. This command does not apply to text entry fields that have a formatting toolbar.
F or **Shift** + **F**	For equation items and text response items, navigates to the previous text entry field. This command does not apply to text entry fields that have a formatting toolbar.
Tab	Moves to the next component on the page. In general, the following test elements are components: • Navigation and tool buttons • Question numbers (and associated prompt text) • Context menus • Response options
Shift + **Tab**	Moves to the previous component on the page.
Return	• Selects response option. • Opens a context menu or list item.
VO + **Right Arrow**	Moves to the next line on the page.
VO + **Left Arrow**	Moves to the previous line on the page.
VO + **A**	Reads everything on the page (from the current point of focus).
VO + **B**	Reads everything from the top of page to the current point of focus.
VO + **U**	Opens Rotor used to interact with objects in web area.
Ctrl	Stops VoiceOver from reading.
Opening and Using Context Menus with VoiceOver Keyboard Commands	
Return	Pressing **Return** when VoiceOver reads “Menu button” will open the context menu. This is the one of two ways to open the context menu when streamlined mode is turned on. Menus can also be opened using **VO** + **Space**.
Down Arrow	Moves the focus to the next option in the menu. VoiceOver will read this option aloud.
Up Arrow	Moves the focus to the previous option in the menu. VoiceOver will read this option aloud.
Space	Selects the menu option currently in focus.
Esc	Closes the context menu without selecting any options.

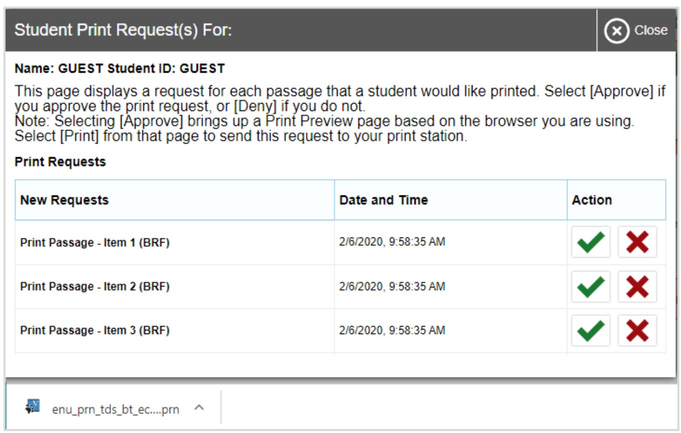
Key	Action
Responding to Items with VoiceOver Keyboard Commands	
Tab	• Students can use the **Tab** key to navigate to the item prompt, which VoiceOver will read aloud. • After VoiceOver reads the prompt aloud, students can press **Tab** again to navigate to the response area. They may need to press **Tab** multiple times depending on the item type and whether any question labels appear for the item. • In the response area for an item, students can press **Tab** to navigate between each answer option, text box, selectable text field, keypad button, or checkbox, depending on the item type.
Shift + **Tab**	Navigates to the previous answer option, text box, selectable text field, keypad button, or check box, depending on the item type.
Up and **Down Arrow Keys**	For edit task choice items, pressing the arrow keys moves the focus between each line of text in the item. After users open an edit task menu by pressing **Space**, the arrow keys can be used to move between the answer options in the list.
VO + **Space** or **Return**	• For multiple choice and multi-select items, pressing either command will select the answer option in focus. • For table match items, pressing either command will mark the checkbox in focus. • For hot text items, pressing either command will select the answer option in focus. ▪ Pressing either command will also choose the selectable text area in focus as the answer option. • For edit task choice items, pressing either command will select an answer option from the drop-down list in the edit menu. • For equation items, pressing either command will select the keypad button in focus.
Option + **7**	For equation items, pressing **Option** + **7** will open a pop-up menu with special characters. Students can use the arrow keys to move between the special characters in the list and then press **VO** + **Space** or **Return** to insert a special character in the response area.

Appendix A. Tips for Embossing Braille Files

This appendix provides tips for approving print requests and embossing braille files.

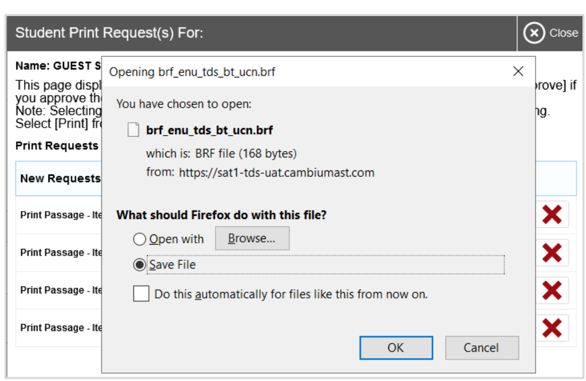
- The auto-emboss setting in TDS automatically queues multiple print requests in advance as students progress through a test. If a student testing with auto-emboss pauses their test before you print all their queued requests, then when the student resumes testing, they must send manual print requests for any unprinted items that were previously in the queue.
- When approving print requests, you can either save a file to your computer and manually open it in the embossing software, or you can open it directly from the browser. The steps for saving and opening braille files depend on which browser you are using:
  - *For Chrome users*: The browser displays the downloaded file at the bottom of the screen (see Figure 22).

Figure 22. Downloading Braille Files in Chrome



    - To open the file directly from the browser, double-click the downloaded file.
    - To save the file to your computer, click the up arrow icon beside the downloaded file and select **Open**.
  - *For Firefox users*: The browser opens a dialog window (see Figure 23).

Figure 23. Downloading Braille Files in Firefox



- To open the file directly from the browser, mark the **Open with** radio button and make sure the appropriate embossing software is selected in the menu (Duxbury Braille Translator for BRF files; Tiger Designer or Tiger Viewer for PRN files). Then click **OK**.

- To save the file to your computer, mark the **Save File** radio button and click **OK**.

– *For Edge users*: The browser shows the downloaded file in the bottom left corner of the screen (see Figure 24).

Figure 24. Downloading Braille files on Chromium-Based Edge

Name: GUEST Student ID: GUEST

This page displays a request for each passage that a student would like printed. Select [Approve] if you approve the print request, or [Deny] if you do not.
Note: Selecting [Approve] brings up a Print Preview page based on the browser you are using. Select [Print] from that page to send this request to your print station.

Print Requests

New Requests	Date and Time	Action
Print Passage - Item 3 (BRF)	2/6/2020, 11:31:20 AM	
Print Passage - Item 2 (BRF)	2/6/2020, 11:31:20 AM	
Print Passage - Item 1 (BRF)	2/6/2020, 11:31:20 AM	

enu_prn_tds_bt_ecn (....p... ...
Open file

- To open the file directly from the browser, click the blue **Open File** link.

- To save the file to your computer, click the three-dot menu icon and select **Show in folder**. This will open File Explorer to the location where the file is saved in your downloads folder.

- If you cannot find a downloaded braille file after approving a print request, you may need to modify your browser's download settings (please contact your local IT support for assistance editing these settings).

- It is highly recommended that you rename downloaded files to correspond with the item number listed on the TA Site's ***Print Request*** window, so you can keep track of printouts when downloading multiple files at once. For example, Item 10 shown in Figure 25 would download with the name *brf_enu_tds_bt_ecn*, but you can simply add *item10* to the beginning of the file name so that it becomes *item10_brf_enu_tds_bt_ecn*.

Figure 25. TA Print Request Window

Print Requests

New Requests	Date and Time	Action
Print Passage - Item 10 (BRF)	10/19/2020, 1:43:59 PM	
Print Passage - Item 11 (BRF)	10/19/2020, 1:44:02 PM	
Print Passage - Item 12 (BRF)	10/19/2020, 1:44:07 PM	

- Always plug the embosser into the same USB port used when it was first set up. Otherwise, the computer may identify the embosser as a new device and require you to set it up again.
- If the **Print** button for a PRN file is grayed out in Tiger Designer or Tiger Viewer, you may need to convert the file in Tiger Designer before embossing it.

User Support

For additional information and assistance in using assistive technology with the Secure Browser, contact the Montana Assessment Helpdesk.

Montana Assessment Helpdesk
(Technical Configuration Support)
Phone Support: 1-877-365-7915
Email Support: mthelpdesk@cambiumasssessment.com

MT Office of Public Instruction Helpdesk
(Policy and Accommodations Approvals)
Phone Support: 1-844-867-2569
Email Support: OPIAssessmentHelpDesk@mt.gov

Please provide the Helpdesk with a detailed description of your problem, as well as the following:

- The brand and version number of the assistive technology being used
- Any error messages and codes that appeared, if applicable
- Operating system and browser information, including version numbers

Appendix 5-C
Test Information Distribution Engine
User Guide

Montana Office of Public Instruction

Test Information Distribution Engine User Guide

2022-2023

Published August 24, 2022

Prepared by Cambium Assessment, Inc.

Table of Contents

Annual Change Log

This log highlights high-level changes within the guide users should be aware of for the 22-23 school year.

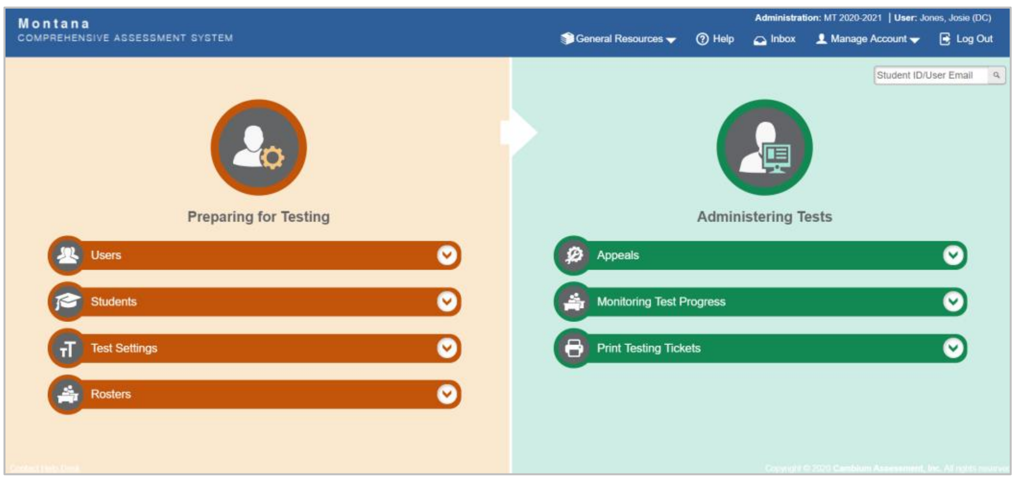
Change	Reason	Key Sections
User Roles updated and re-labeled. All school-level users will need to be re-added by district-level users for 22-23.	Aligns with the OPI's Data Privacy and Security Policies in the MontCAS Test Security Manual.	How to Activate Your Account & Log in to & out of TIDE How District-level Users Set Up User Accounts in TIDE User Role Permissions
Student Test Settings have been updated to better integrate with Infinite Campus and create consistency across programs.	Updated UAAGs for Smarter Balanced, MSA, and Alternate MSA.	Fields in the Test Settings and Tools Panels Valid Values for Tool Names in the Test Settings Upload File

Introduction to TIDE

This user guide provides instructions on how to use TIDE.

At its core, Test Information Distribution Engine (TIDE) is a registration system for users who will access Cambium Assessment, Inc. (CAI) systems and students who will take the OPI's state summative assessments (i.e., Smarter Balanced, Montana Science Assessment (MSA), and the Alternate MSA. Users of all CAI systems must be added to TIDE before they can access any CAI system. Students must be added to TIDE before they can test in the Test Delivery System (TDS). Rosters must be added in TIDE so Reporting can display scores at the classroom-, school-, district-, and the state-level. During testing, TIDE users can print test tickets, manage appeal requests, and monitor testing progress.

Figure 1. TIDE Dashboard



You can use TIDE to perform the following tasks:

- You can add new **users** or modify existing **user accounts** in TIDE so district and school personnel can access TIDE and other CAI systems. Users must be registered in TIDE to access other CAI systems.
- You can modify **student test settings** so students can take the correct tests with the correct test settings at the correct time. Students must be registered in TIDE to test in TDS.
- You can add new rosters or modify existing rosters. Rosters represent classes or other groups of students. After testing, TIDE sends rosters to Reporting so that system can display scores at the classroom, school, district, and state levels.
- You can print hard-copy **test tickets** that include a student's username so the student can log in to a test.
- You can add new appeal requests or modify existing appeal requests if a test must be retaken or reopened.
- You can view your district's or school's progress in **starting and completing tests** and **participation rate**.

TIDE divides tasks by user role. Users with higher roles will have access to more tasks in TIDE than users with lower roles. District-level users have access to the most tasks, followed by school-level users, Teachers/Test Administrators (TA Role), and Proctors (PR Role). The structure of this guide is based on user role. It includes the following sections:

- How to Activate Your Account & Log in to & out of TIDE
- How District-level Users Perform Tasks in TIDE
- How School-level Users Perform Tasks in TIDE
- How Teachers/Test Administrators (TA Role) and Proctors (PR Role) Proctors Perform Tasks in TIDE

There is also an Appendix with additional information and instructions.

Three Things All TIDE Users Must Know How To Do

Records for users, students, and rosters must be added to TIDE and kept up to date for the testing process to flow properly. Users not added to TIDE will not have access to any CAI systems. Students not added to TIDE will not be able to test. If a student is not present in TIDE, check with your System Testing Coordinator (STC) or Building Coordinator (BC) and ensure the student information is accurate in the district's student information system (SIS) that syncs to AIM/Infinite Campus.

How are Students Registered in TIDE to Test?

Montana's procedures for determining which students are eligible to test uses the current fiscal year school records that begins on July 1 and ends on June 30 of each year (see §20-1-301 MCA). The OPI expects all students enrolled in accredited schools with "primary" educational service type records in the State Student Information System (i.e., Achievement in Montana (AIM)/Infinite Campus System) to test. Only student records in AIM/Infinite Campus for primary enrolled students are provided to the CAI systems.

The accountability and state assessment participation expectations are set from several OPI determined collections published in the OPI's AIM/Infinite Campus Collection Schedule. Any student enrolled on the test count date and during the OPI's Published Test Windows are expected to participate in the state summative assessments. If the student is enrolled in an accredited school in AIM/Infinite Campus with a primary enrollment and also included in MAEFAIRS for purposes of determining the Average Number Belonging (ANB), then the OPI expects the student to participate in state assessments. For more details on the school district's obligations for state testing, accountability, and reporting, read the OPI's ESSA State Plan.

Rosters not added to TIDE will not be available in Reporting and you will not be able to view your students' test results by class or by other meaningful groupings. The process for adding and modifying records in TIDE is user-friendly because it's basically the same no matter your user role or which type of record you want to add.

All TIDE users must be familiar with the following actions, as they are the same for Users, Rosters, and Appeal Requests:

- You can **add** new records one at a time.
- You can **view or modify** existing records one at a time.
- You can **upload** multiple new records or modify multiple existing records in the same file.

How to add records one at a time

1. Start at the dashboard that appears when you first log in to TIDE, select the task for which you want to add a new record, and select **Add**.
2. On the page that appears, fill out the information, verify its accuracy, and select **Save**.

Figure 2. Add User Page

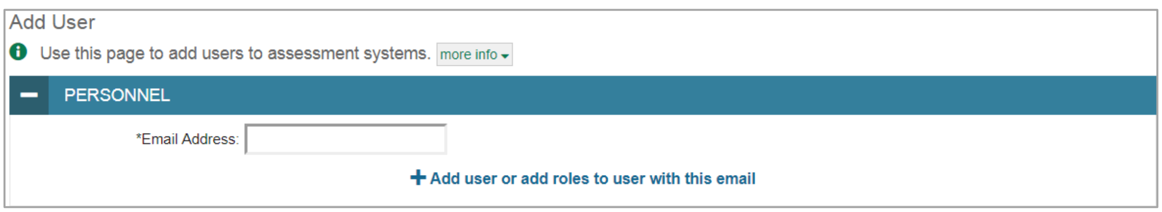
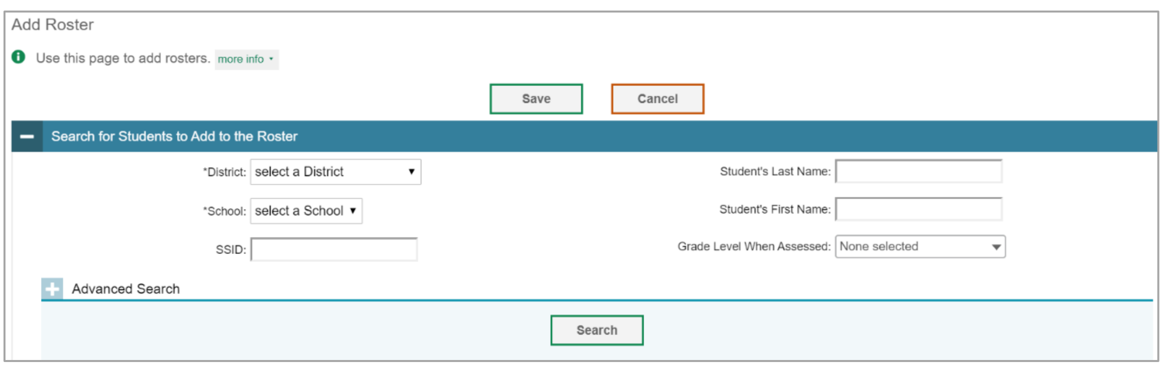


Figure 3. Add Roster Page

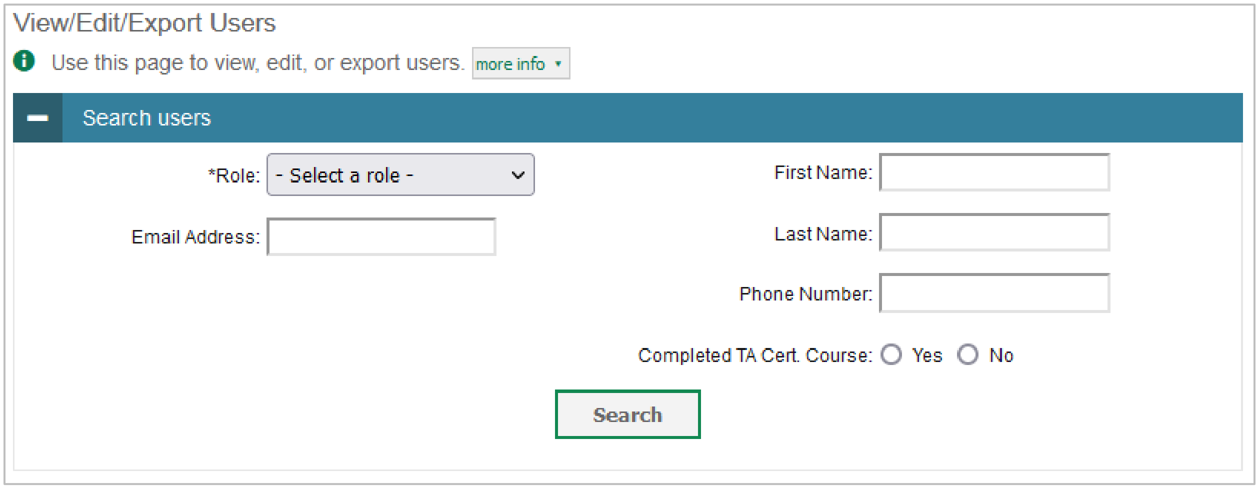


How to modify existing records one at a time

You can view and edit existing records for users in the school system (such as BC, TA, PR) one at a time or multiple existing records all at once through file export. Student records cannot be modified. If a record's information changes after you've added the record to TIDE, you must edit the record to match the most up to date information. You can also delete records from TIDE.

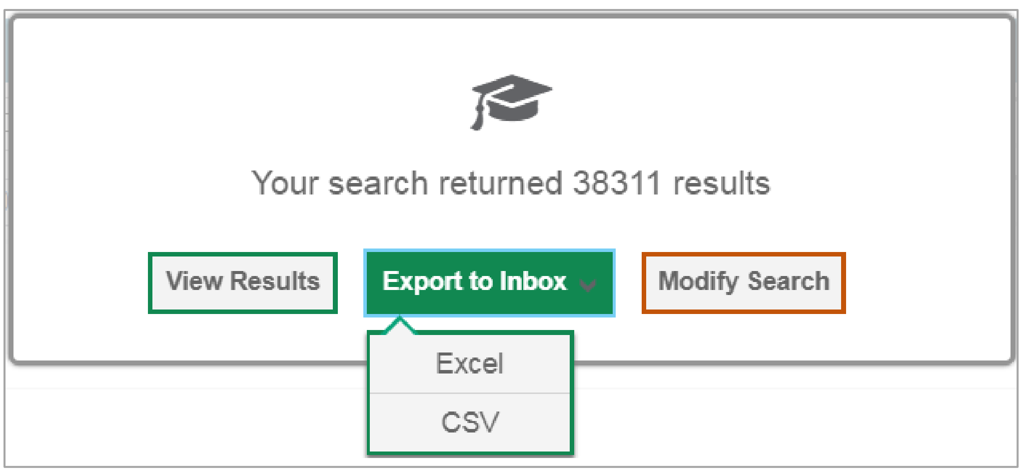
1. Begin by searching for the record you want to modify. Start at the dashboard that appears when you first log in to TIDE, select the task for which you want to search for records, and select **View/Edit/Export**. Fill out the form that appears and select **Search**.

Figure 4. View/Edit/Export User Page



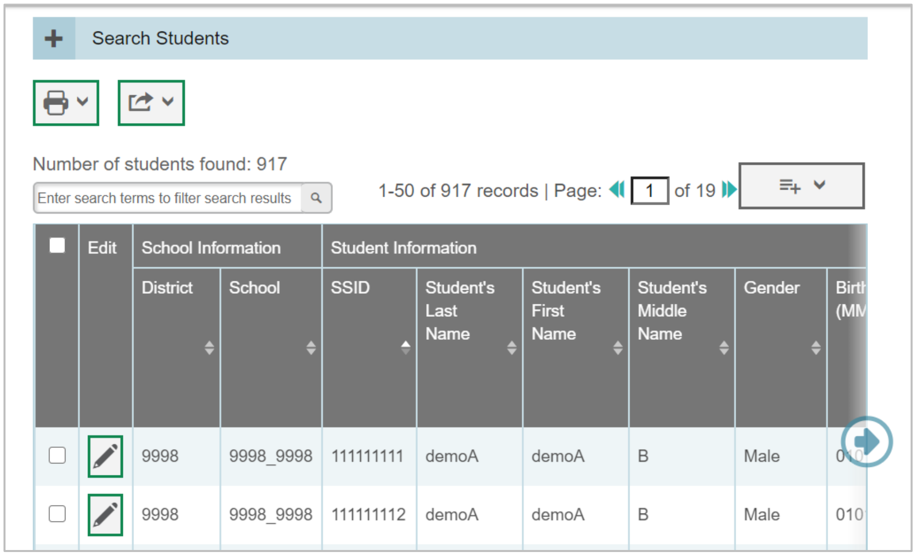
2. A pop-up window appears, allowing you to view or export search results or modify your search. To view and edit search results, select **View Results**. To export all search results to the inbox from the pop-up window, select **Export to Inbox** and then select either **Excel** or **CSV**. The search results will be exported to your inbox and you will return to the search form.

Figure 5. Search Results



3. If you select **View Results**, the search results will appear in a table. To edit individual records, select the edit button by the record you want. To export records, mark the checkbox by that record and select .

Figure 6. View/Edit/Export Student Page with Search Results

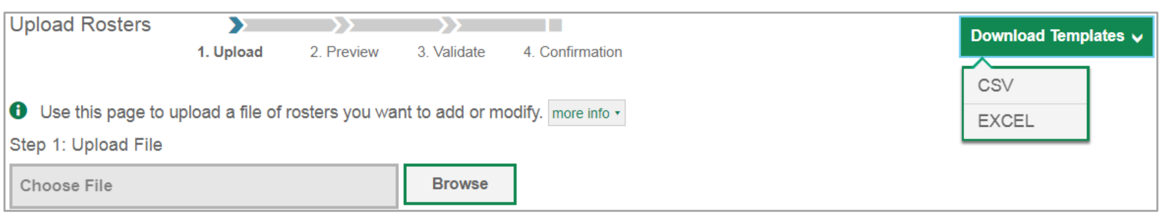


How to add or modify multiple records at once

Rather than adding or modifying records one at a time, you may want to add or modify multiple records all at once. File upload allows you to do this. Records not previously set up in TIDE will be added to TIDE through file upload. Records already set up in TIDE will be modified with the updated content from the upload. To upload records, you must be familiar with spreadsheet applications and/or comma-separated value (CSV) files.

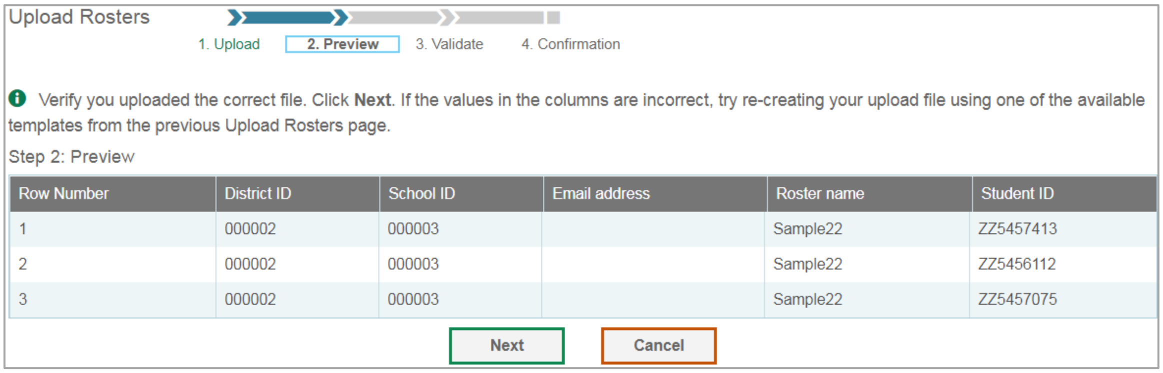
1. Start at the dashboard that appears when you first log in to TIDE, select the task for which you'd like to upload records, and select **Upload**. An upload screen will appear where you can download a template file.

Figure 7. Upload Rosters Page



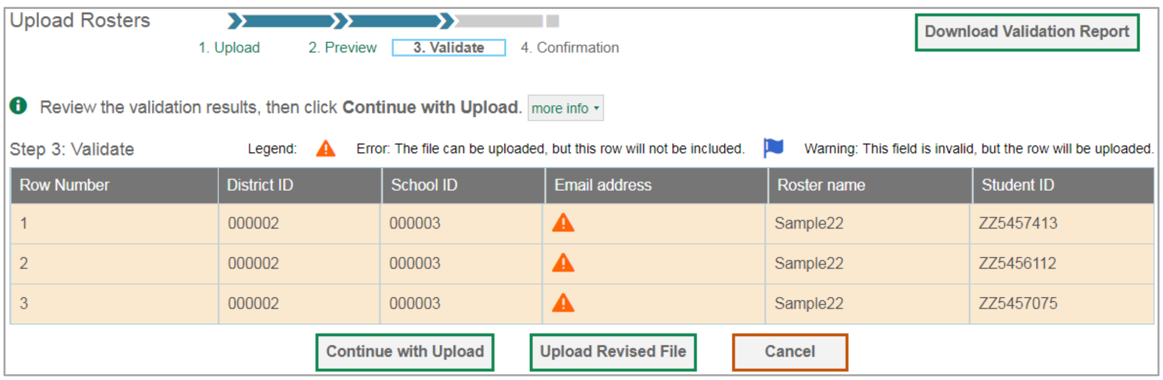
2. Once you've downloaded and filled out the template file, return to the upload screen, select **Browse**, locate the file on your computer, and upload it to TIDE. Select **Next**. The upload preview screen appears.

Figure 8. Upload Rosters Preview Page



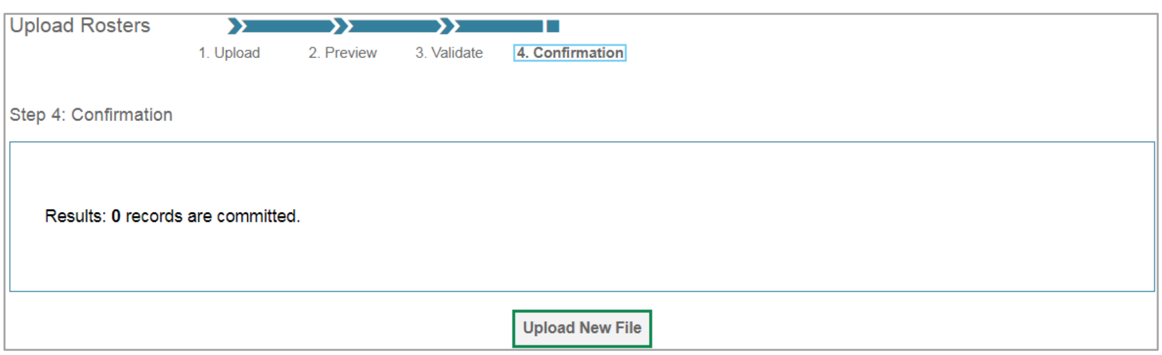
3. Once you've verified the information on the preview screen, select **Next** again. The validation screen appears.

Figure 9. Upload Rosters Validation Page



4. The validation screen shows errors or warnings associated with your uploaded file. To continue with the upload despite these errors or warnings, select **Continue with Upload**. The confirmation screen appears. To revise the file before uploading, select **Upload Revised File**. To upload a new file from the confirmation screen, select **Upload New File**.

Figure 10. Upload Rosters Confirmation Page



How to Activate Your Account & Log in to & out of TIDE

Your Authorized Representative (AR) or System Test Coordinator (STC) for the school district creates your account, and then TIDE sends you an activation email. This email contains a link that takes you to the ***Reset Your Password*** page in TIDE where you can set up your password for logging in to TIDE and other applicable CAI systems. This link expires 15 minutes after the email was sent. If you do not set up your password within 15 minutes, you need to request for a new link as described in the section "Password Information" in the appendix.

If you do not receive an activation email, check your spam folder. Emails are sent from DoNotReply@cambiumassessment.com, so you may need to add this address to your contact list.

At the beginning of a new school year, your TIDE password and security details will be automatically reset. You will receive an email from DoNotReply@cambiumassessment.com to notify you of this occurrence and to alert you that you will not be able to log in to TIDE or any other system until you reactivate your account for the new school year. Follow the instructions in the section "How to reactivate your account" below to reactivate your account for the new school year.

How are User Permissions Determined?

All school-level user roles in the Montana Test Portal System [TIDE] are annually refreshed to match the "User Roles and Access Document" in the system. Only authorized users are permitted to access and manage these system accounts. Specific usernames and passwords allow access to confidential school and student data. All educators involved in the administration of state assessments are responsible for maintaining the privacy and security of student records specified under the federal Family Educational Rights and Privacy Act (FERPA) and by state law. As referenced in the MontCAS Test Security Manual, state assessments may only be administered to students by licensed professionals who have been trained, are familiar with standardized testing procedures, and are employed by the school system.

This school year, TIDE has undergone several system updates to better align with the OPI's policies and procedures for data privacy and needs across state assessment programs. The process of refreshing existing user roles in TIDE is a standard maintenance practice to enforce policies on security and privacy for access to student data.

How to activate your account

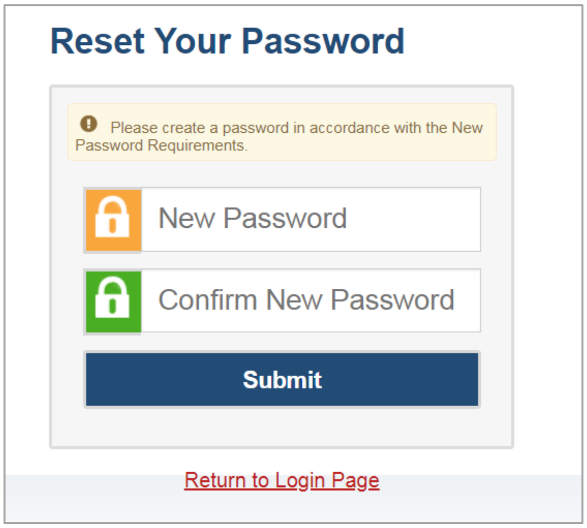
1. Select the link in the activation email. The ***Reset Your Password*** page appears (see Figure 11).

2. In the *New Password* and *Confirm New Password* fields, enter a new password. The password must be at least eight characters long and must include at least one lowercase alphabetic character, one uppercase alphabetic character, one number, and one special character (e.g., %, #, or !).

3. Select **Submit**.

Account activation is complete. You can proceed to TIDE by selecting the **TIDE** card (see Figure 13) in the portal page.

Figure 11. Reset Your Password Page



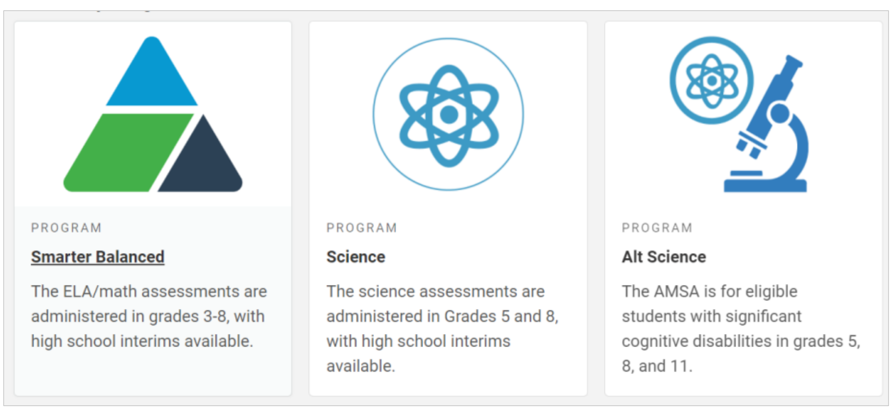
How to reactivate your account at the beginning of the school year

At the beginning of a new school year, your TIDE password and security details may be automatically reset. If so, you will receive an email from DoNotReply@cambiumassessment.com to notify you of this occurrence and to alert you that you will not be able to log in to TIDE or any other system until you reactivate your account for the new school year.

1. Navigate to the Montana Assessment Portal (http://mt.portal.cambiumast.com).

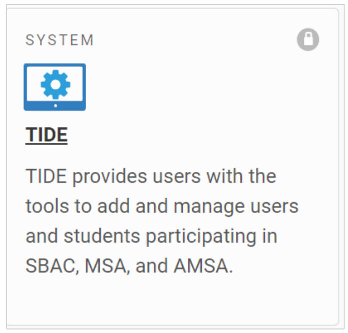
2. Select the assessment card you will be administering.

Figure 12. Assessment Cards on Portal



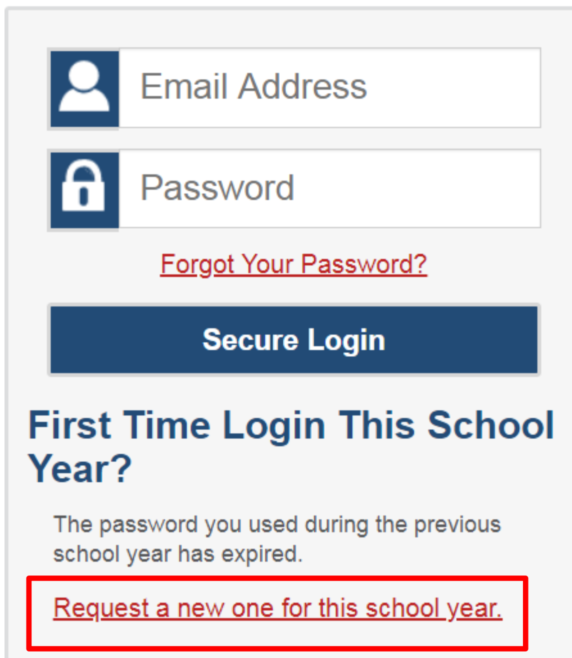
3. Select **TIDE** (see Figure 13). The ***Login*** page appears (see Figure 14).

Figure 13. TIDE Card



4. Select **Request a new one for this school year**. The ***Reset Your Password: Find Account*** page appears (see Figure 15).

Figure 14 Login Page



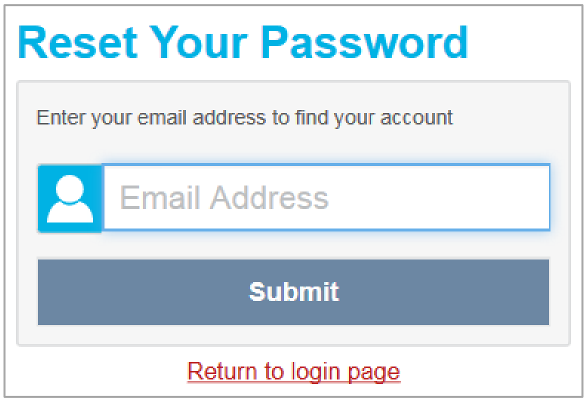
5. Enter your TIDE email address and select **Submit**. TIDE sends you an email containing a link to reset your password.

6. Select the link in the activation email. The ***Reset Your Password*** page appears (see Figure 11).

7. In the *New Password* and *Confirm New Password* fields, enter a new password. The password must be at least eight characters long and must include at least one lowercase alphabetic character, one uppercase alphabetic character, one number, and one special character (e.g., %, #, or !).

8. Select **Submit**.

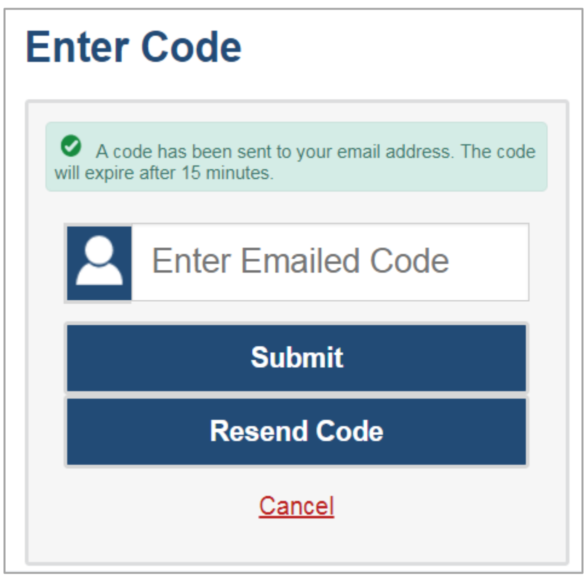
Figure 15. Reset Your Password Page



During the reactivation process, you will be taken to the ***Enter Code*** (see Figure 16) page and asked to provide the authentication code sent to your email.

- In the *Enter Emailed Code* field, enter the emailed code and select **Submit**.
- You must enter the code within fifteen minutes of the email being sent. If your code expires, you can request for a new code by selecting **Resend Code** on the ***Enter Code*** page.

Figure 16. Enter Code Page



How to log in to TIDE

Do not share your login information with anyone. All Montana Assessment systems provide access to student information, which must be protected in accordance with federal privacy laws.

1. Navigate to the Montana Assessment Portal (http://mt.portal.cambiumast.com).

2. Select the assessment card you will be administering (See Figure 12).

3. Select **TIDE** (see Figure 13). The ***Login*** page appears (see Figure 14).

4. On the ***Login*** page, enter the email address and password you use to access all CAI systems.

5. Select **Secure Login**.

    a. If you have not logged in using this browser before, or if you have cleared your browser cache, the ***Enter Code*** page appears (see Figure 16) and an email is sent to your address. This applies every time you access TIDE with a new browser. The email contains an authentication code, which you must use within fifteen minutes of the email being sent.

        i. In the *Enter Emailed Code* field, enter the emailed code. If the code has expired, Select **Resend Code** to request a new code.

        ii. Select **Submit**.

    b. If you are logging in for the first time this school year, the ***Test Security Agreement*** page appears. You should review the terms on this page carefully in accordance with Montana Office of Public Instruction state policy. Click **Accept** to acknowledge and proceed.

What is the Purpose of the Test Security Agreement

The Building Coordinator (BC) is a person is typically assigned by the Authorized Representative (AR) and/or System Test Coordinator (STC) as a licensed non-instructional person such as principals, vice principals, counselors, or other staff members. This person coordinates state assessments in the school building. All staff with TIDE Testing Portal accounts will be required to sign an electronic Test Security Agreement (TSA). The Authorized Representative (AR) and Building Coordinator (BC) Roles and Responsibilities for Test Security outline responsibilities the school and district level roles must be aware of for proper test security.

The OPI expects that school districts will annually monitor TIDE for completed TSAs and Test Administrator (TA) training/certification requirement before administering any assessments. and contact the OPI for technical assistance as needed. For more information on this topic, read the MontCAS Test Security Manual or contact the OPI Assessment Help Desk at opiassessmenthelpdesk@mt.gov or 1-844-867-2569.

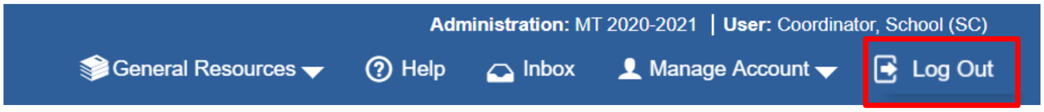
The ***Dashboard*** for your user role appears. Depending on your user role, TIDE may prompt you to select a role, district, or school to complete the login.

Working with TIDE in more than one browser tab or window may result in changes in one tab overwriting changes made in another tab. Do not have more than one TIDE browser tab or window open at one time.

How to log out of TIDE

- In the TIDE banner (see Figure 17), select **Log Out**.

Figure 17. Log Out



Logging out of TIDE logs you out of most Montana Assessment systems housed within the Montana Testing Portal. However, you will not be logged out of the TA Interface in order to prevent the accidental interruption of active test sessions.

How District-level Users Perform Tasks in TIDE

District-level users (e.g., AR and STC) can perform most of the tasks available in TIDE. Some of these tasks must be performed before testing begins, some must be performed during testing, and some must be performed after testing.

How District-level Users Perform Tasks in TIDE Before Testing Begins

Before testing begins, district-level users must perform the following tasks in TIDE:

- Set up **user accounts** for school-level users so they can log in to TIDE and other CAI systems. If user accounts are not set up before testing begins, those users will not be able to access any CAI systems.
- Update **student test settings** to ensure correct tests are assigned. If student accounts and correct test settings are not present in TIDE before testing begins, those students will not be able to test. Contact your district's System Test Coordinator (STC) if a student is not present in TIDE.
- Set up **rosters** so Reporting can display scores at the classroom, school, district, and state levels.

How District-level Users Set Up User Accounts in TIDE

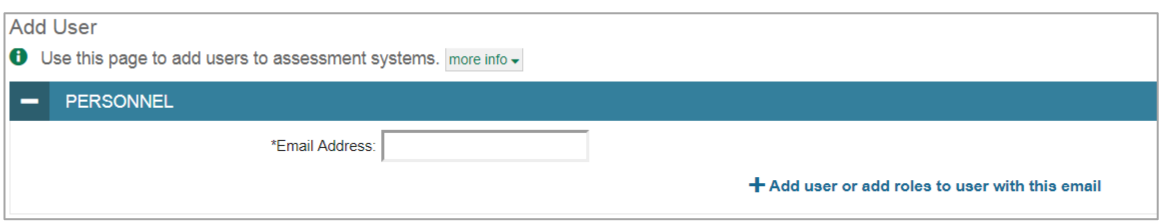
District-level users must set up user accounts for school-level users to sign in to TIDE and other CAI systems. If these users don't have accounts set up in TIDE, they will not be able to access any CAI systems.

How district-level users add new user accounts one at a time

You can add users to TIDE one at time. To learn more about adding records to TIDE one at a time, see the section "How to add records one at a time" in the Introduction.

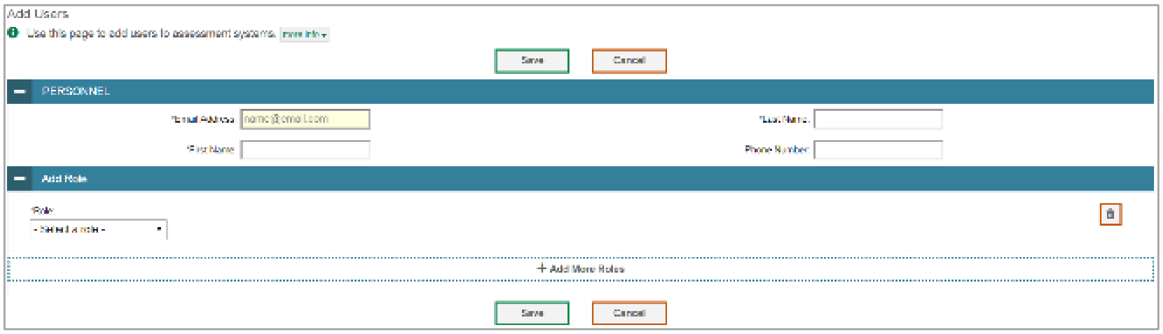
1. From the **Users** task menu, select **Add Users**. The ***Add Users*** page appears.

Figure 18. Add User Page



2. In the *Email Address* field, enter the new user's email address and select **+Add user or add roles to use with this email**. Additional fields appear.
3. Enter the new user's first and last names in the required fields and other details in the optional fields.

Figure 19. Add User Page



4. From the **Role** drop-down, select a role. From the drop-downs that appear, select a state, district, and school, if applicable.

5. *Optional:* To add multiple roles, select **+Add More Roles** and repeat step 4.

6. *Optional:* To delete a role, select [delete icon] next to that role.

7. Select **Save**. In the affirmation dialog box, select **Continue** to return to the ***Add Users*** page. TIDE adds the account and sends the new user an activation email from DoNotReply@cambiumassessment.com.

How district-level users modify existing user accounts one a time

You can view and modify existing user accounts one at a time or multiple existing user accounts all at once through file export. If a user's information changes after you've added the user to TIDE, you must edit the user account to match the most up to date information. If the user's account does not include the most up to date information, the user may not be able to access other CAI systems or features within those systems. You can also delete users from TIDE.

1. From the **Users** task menu, select **View/Edit/Export Users**. The ***View/Edit/Export Users*** page appears.

2. Retrieve the individual user account you want to view, edit, export, or delete by following the procedure in the section "How to modify existing records one at a time" in the Introduction.

3. In the list of retrieved user accounts, select [edit icon] for the user whose account you want to view or edit.

4. If your role allows it, modify the user's details as required, using the table "Fields in the View/Edit Users [User's Name] Page" in the appendix as a reference.

5. *Optional:* To add more roles for this user, select **+Add More Roles** and then follow the steps as described in the section on adding individual users.

6. *Optional:* To delete a role, select [delete icon] next to that role. You can also delete the user's entire account from the search results table.

7. Select **Save**.

8. In the affirmation dialog box, select **Continue** to return to the list of user accounts.

How district-level users add or modify multiple user accounts all at once

You can also add or modify multiple user accounts all at once through file upload as shown in the section "How to add or modify multiple records at once" in the Introduction.

1. From the **Users** task menu, select **Upload Users**. The ***Upload Users*** page appears.

2. Following the instructions in the section "How to add or modify multiple records at once" in the Introduction and using the table "Columns in the User Upload File" in the appendix as a reference, fill out the template and upload it to TIDE. Users who have not previously been set up in TIDE will be added in TIDE. Users who already have accounts set up in TIDE will have their accounts modified with the updated content from the upload.

How district-level users upload user attributes

You can set up attributes for multiple users through file uploads. This task requires familiarity with composing comma-separated value (CSV) files or working with Microsoft Excel.

1. From the **Users** task menu, select **Upload User Attributes**. The ***Upload User Attributes*** page appears.

2. Following the instructions in the section "How to add or modify multiple records at once" in the Introduction and using the table "Columns in the User Upload File" in the appendix as a reference, fill out the Attribute template and upload it to TIDE.

How District-level Users Update Students for Testing

How district-level users specify student accommodations and test settings

A student's test settings include the available accommodations and designated supports, such as text-to-speech or color choices. This section explains how to edit student test settings via an online form or a file upload.

1. From the **Test Settings and Tools** task menu on the TIDE dashboard, select **View/Edit/Export Test Settings and Tools**. The ***View/Edit/Export Test Settings and Tools*** page appears.

2. Retrieve the student accounts whose settings and tools you want to view or edit by following the procedure in the section "How district-level users modify existing student accounts one at a time."

3. In the list of retrieved students, select  for the student whose test settings and tools you want to edit. The ***View/Edit Students: [Student's Name]*** form appears.

How district-level users upload student accommodations and test tools

If you have many students for whom you need to apply test settings, it may be easier to perform those transactions through file uploads. This task requires familiarity with composing comma-separated value (CSV) files or working with Microsoft Excel.

1. From the **Test Settings and Tools** task menu on the TIDE dashboard, select **Upload Test Settings and Tools**. The ***Upload Test Settings and Tools*** page appears.

2. Following the instructions in the section "How to add or modify multiple records at once" in the Introduction and using the table "Columns in the Test Settings Upload File" in the appendix as a reference, fill out the Test Settings template and upload it to TIDE.

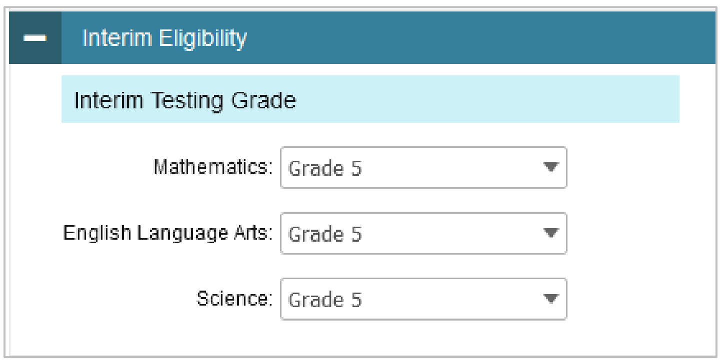
How district-level users set student eligibilities for interim assessments

You can add eligibility rules for students so they can access the proper interim tests.

1. From the **Students** task menu on the TIDE dashboard, select **View/Edit/Export Students**. Navigate to the ***Interim Eligibility*** section on the ***View/Edit Students: [Student's Name]*** form (see Figure 20).

2. Select the grade level for testing.

Figure 20. Drop-down Lists on the Interim Eligibility Section



How district-level users upload interim grades

You can set up interim grades for multiple students through file uploads. This task requires familiarity with composing comma-separated value (CSV) files or working with Microsoft Excel.

1. From the **Students** task menu on the TIDE dashboard, select **Upload Interims**. The ***Upload Interims*** page appears.

2. Following the instructions in the section "How to add or modify multiple records at once" in the Introduction and using the table "Columns in the Interim Grades Upload File" in the appendix as a reference, fill out the Interim Grade template and upload it to TIDE.

About the Interim Grades Upload File

If the upload file includes two rows specifying different grades for the same student and subject, then both grades will be set up as interim grades for the student's subject.

If the upload file includes two rows for the same student and subject and the second row has a value "None", then all interim grades established for the student's subject up to that point will be removed.

How District-level Users Manage Rosters

Rosters are groups of students associated with a teacher in a particular school. Rosters typically represent entire classrooms in lower grades, or individual classroom periods in upper grades. Rosters can also represent special courses offered to groups of students.

The rosters you create in TIDE are available in Reporting. Reporting can aggregate test scores at these roster levels. You can also use rosters to print test tickets containing students' login information to start taking a test.

Since Teachers/Test Administrators (TA Role) are responsible for the growth and development of student's skill-sets, such as reading, writing, research, communication, and problem solving, it is important for a teacher to be able to analyze his students' performance data and adjust his teaching strategies accordingly. For a teacher to be able to see his students' performance data, the students must be included in a roster associated with the teacher. Hence, rosters need to be created for all Teachers/Test Administrators (TA Role) who are responsible for teaching an academic subject, such as Reading/Literacy, Mathematics, and Science.

When creating rosters, it is recommended to follow the guidelines below:

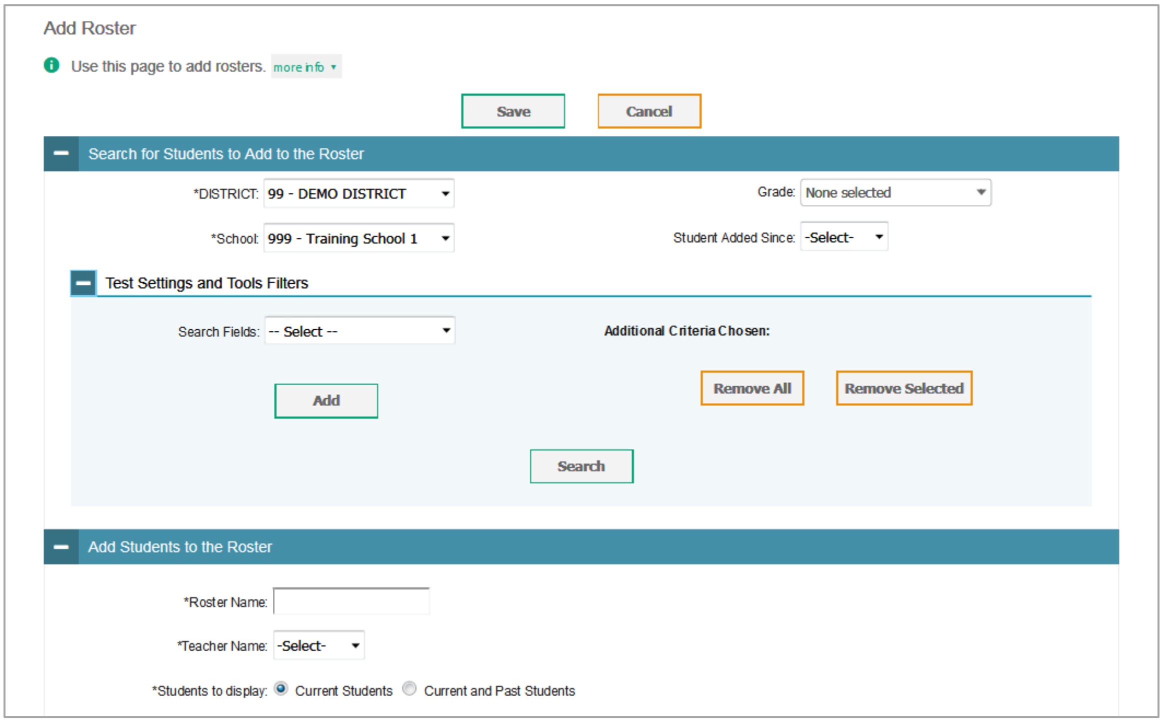
- Rosters should ideally include about 25–30 students. If a roster is too large or too small, it may affect the credibility and usefulness of the data.
- One or more rosters may need to be created depending on the subjects taught by a teacher. For example, if a group of Grade 3 students have the same teacher for Reading, Mathematics, and Science, then separate rosters do not need to be created for each subject. However, if different Teachers/Test Administrators (TA Role) are responsible for teaching different subjects then separate rosters need to be created for each teacher and subject.
- When naming rosters, a clear and consistent naming convention should be used that indicates the grade, class name, teacher, period as applicable. For example, an elementary school roster may be named 'Gr3Jones17-18' and a secondary school roster may be named 'AikenPeriod3Eng9A17-18'.

You can only create rosters from students associated with your school or district.

How district-level users add new rosters one at a time

1. From the **Rosters** task menu on the TIDE dashboard, select **Add Roster**. The ***Add Roster*** form appears (see Figure 21).
2. In the *Search for Students to Add to the Roster* panel, search for students by filling out the search criteria and selecting **Search**.

Figure 21. Add Roster Form



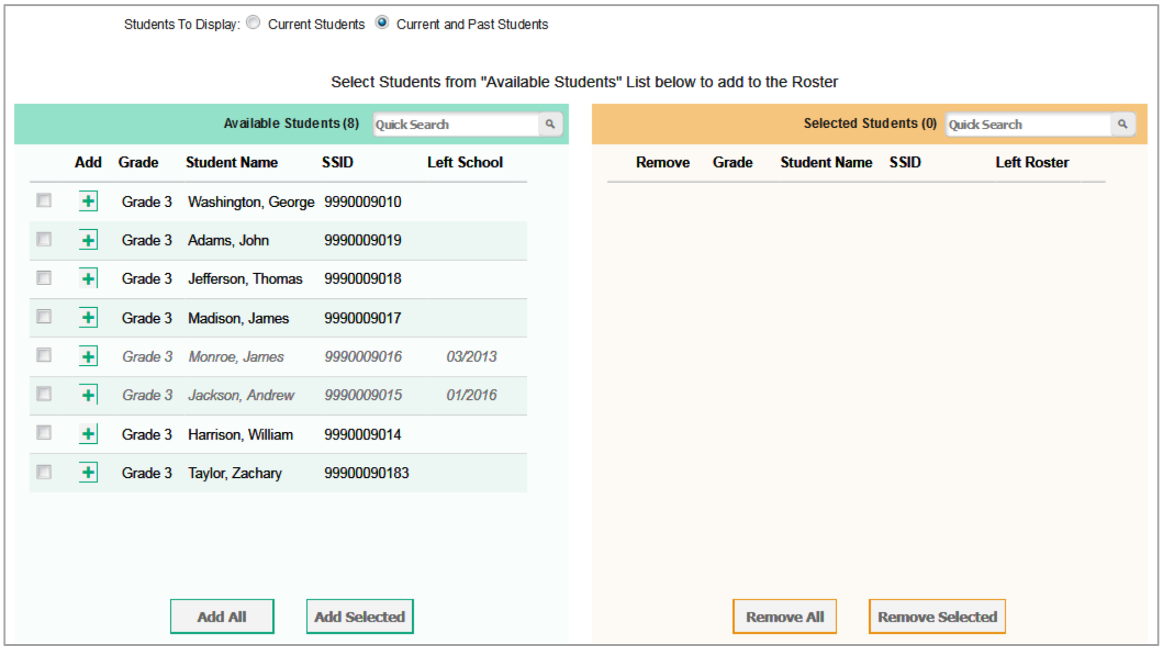
3. In the *Add/Remove Students to the Roster* panel (see Figure 22), do the following:

   a. In the *Roster Name* field, enter the roster name.

   b. From the *Teacher Name* drop-down list, select a teacher or school personnel associated with the roster.

   c. From the *Students to display* field, select the students you wish to view in the *Available Students* list. The two options are:

Figure 22. Detail of Add/Remove Students to Roster Panel



- **Current Students**: Displays students who match your search criteria and are currently associated with the school.

- **Current and Past Students**: Displays all the students who match your search criteria from the current year even if they are no longer associated with the school. For example, if a Grade 3 student has left the school and you search for Grade 3 students with the *Students to display* field set to **Current and Past Students**, the student who has left the school will also be displayed.

When viewing current and past students from the selected year, students who are no longer associated with your school will display the date on which they left the school. You can still add these students to your roster, if desired.

d. To add students, in the list of available students do one of the following:

- To move one student to the roster, select ➕ for that student.

- To move all the students in the *Available Students* list to the roster, select **Add All**.

- To move selected students to the roster, mark the checkboxes for the students you want to add, then select **Add Selected**.

e. To remove students, do one of the following in the list of students in the roster:

- To remove one student from the roster, select ✖ for the student.

- To remove all the students from the roster, select **Remove All**.

- To remove selected students from the roster, mark the checkboxes for the students you want to remove, then select **Remove Selected**.

2. Select **Save**, and in the affirmation dialog box, select **Continue**.

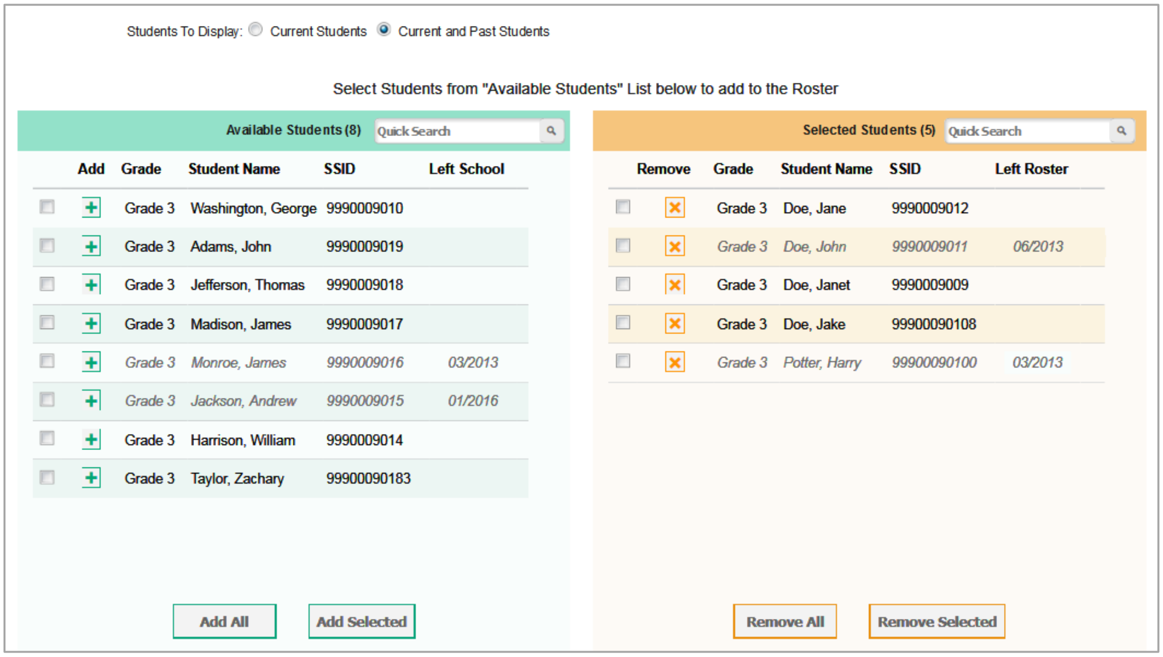
How district-level users modify existing rosters one at a time

You can modify existing rosters by performing the following steps:

1. From the **Rosters** task menu on the TIDE dashboard, select **View/Edit/Export Roster**. The ***View/Edit/Export Roster*** page appears.
2. Retrieve the roster record you want to view or edit by following the procedure in the section "How to modify existing records one at a time" in the Introduction.
3. In the list of retrieved rosters, select for the roster whose details you want to view. The ***View/Edit Roster*** form appears. This form is similar to the form used to add rosters (see Figure 21).
4. In the *Search for Students to Add to the Roster* panel, search for students by following the procedure in the section "How to modify existing records one at a time" in the Introduction.
5. In the *Add/Remove Students to the Roster* panel (see Figure 22), do the following:
    a. In the *Roster Name* field, enter the roster name.
    b. From the *Teacher Name* drop-down list, select a teacher or school personnel associated with the roster.
    c. From the *Students to display* field, select the students you wish to view in the *Available Students* and *Selected Students* lists. The two options are:
        - **Current Students**: Displays students who match your search criteria and are currently associated with the school and roster. The *Available Students* list displays students who are currently associated with your school and the *Selected Students* list displays students who are currently associated with the roster.
        - **Current and Past Students**: Displays all the students who match your search criteria from the current year even if they are no longer associated with the school or the roster. If a student has been removed from the roster, the date on which he was removed from the roster is displayed in the *Selected Students* list. If the student who has been removed from the roster is still associated with the school, he is listed in the *Available Students* list as a regular student. However, if he has left the school then his record will appear in the *Available Students* list with the date he left the school.
    d. To add students, from the list of available students, do one of the following:
        - To move one student to the roster, select for that student.
        - To move all the students in the *Available Students* list to the roster, select **Add All**.

- To move selected students to the roster, mark the checkboxes for the students you want to add, then select **Add Selected**.

Figure 23. Modifying a Roster: Current and Past Students



e. To remove students, do one of the following in the list of students in the roster:

   - To remove one student from the roster, select ✖ for the student.
   - To remove all the students from the roster, select **Remove All**.
   - To remove selected students from the roster, mark the checkboxes for the students you want to remove, then select **Remove Selected**.

6. Select **Save**, and in the affirmation dialog box select **Continue**.

How district-level users add or modify multiple rosters all at once

If you have many rosters to add or modify, you can do so through file upload as shown in the section "How to add or modify multiple records at once" in the Introduction

1. From the **Rosters** task menu on the TIDE dashboard, select **Upload Rosters**. The ***Upload Rosters*** page appears.
2. Following the instructions in the section "How to add or modify multiple records at once" in the Introduction and using the table "Columns in the Roster Upload File" in the appendix as a reference, fill out the Roster template and upload it to TIDE.

How District-level Users Use TIDE during Test Administration

During testing, district-level users can perform the following tasks in TIDE:

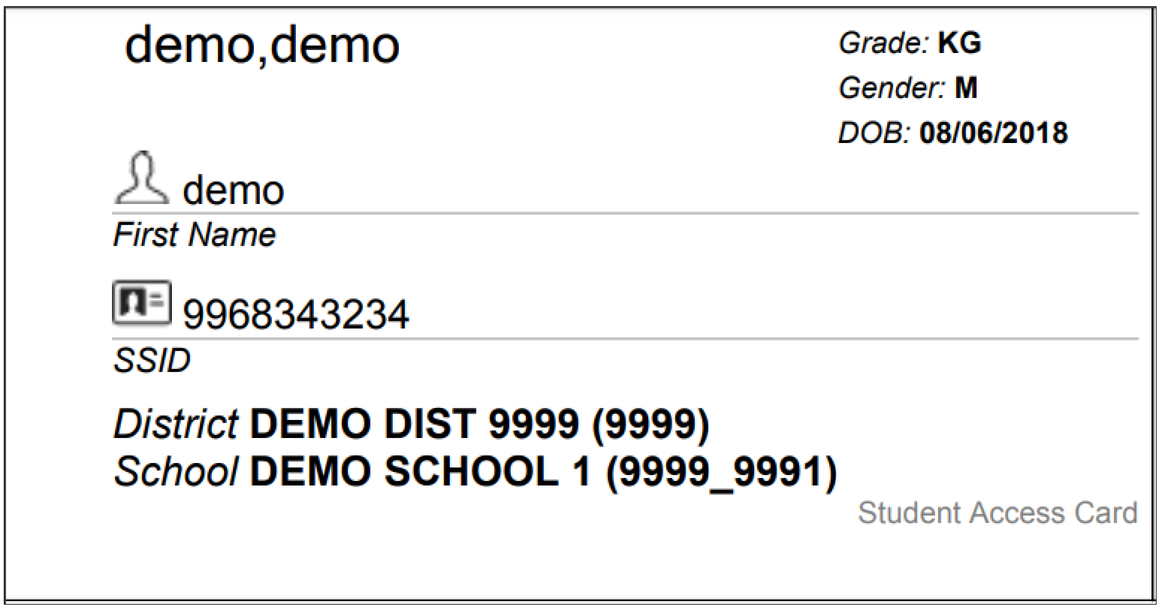
- Print **test tickets** to help students log in to tests.
- Add, modify, and upload appeal requests.
- View reports of students' current test statuses, test completion rates, and test status codes.

How District-level Users Print Test Tickets

A test ticket is a hard-copy form that includes a student's username for logging in to a test.

TIDE generates the test tickets as PDF files that you download with your browser.

Figure 24. Sample Test Ticket



demo,demo

Grade: **KG**
Gender: **M**
DOB: **08/06/2018**

demo
First Name

9968343234
SSID

District **DEMO DIST 9999 (9999)**
School **DEMO SCHOOL 1 (9999_9991)**

Student Access Card

How district-level users print test tickets from student lists

1. From the **Print Test Tickets** task menu on the TIDE dashboard, select **Print from Student List**. The ***Print Test Tickets from Student List*** page appears.
2. Retrieve the students for whom you want to print test tickets by filling out the search criteria and selecting **Search**.
1. Select the column headings to sort the retrieved students in the order you want the test tickets printed.

2. Specify the students for whom test tickets need to be printed:

   - To print test tickets for specific students, mark the checkboxes for the students you want to print.
   - To print test tickets for all students listed on the page, mark the checkbox at the top of the table.
   - To print test tickets for all retrieved students, no additional action is necessary. The option to print all retrieved records is available by default.

3. Select  and then select the appropriate action:

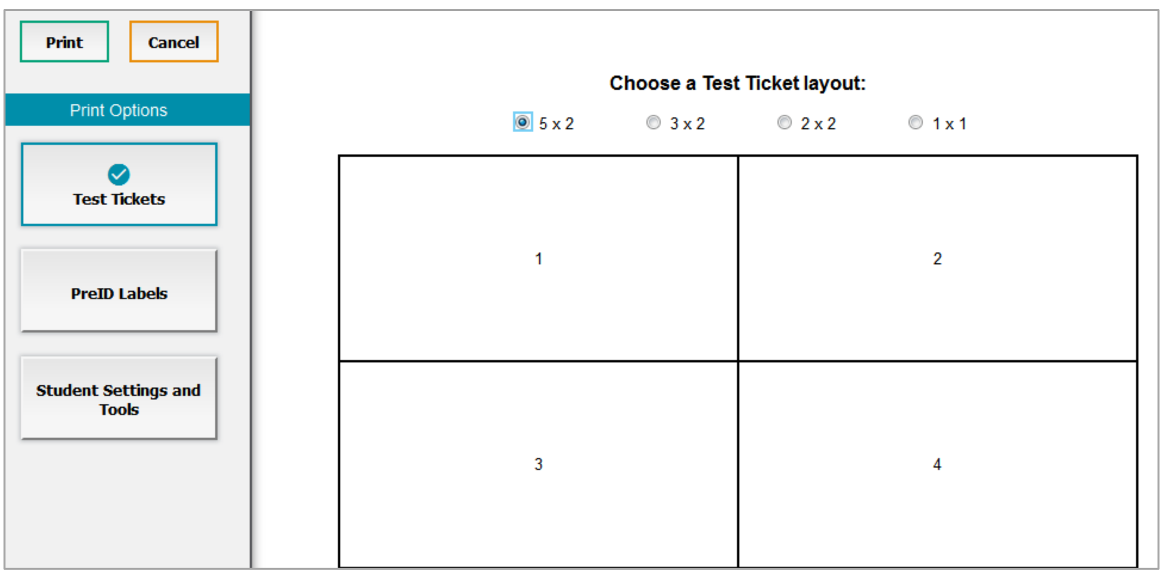
   - To print test tickets for selected students, select **My Selected Test Tickets**.
   - To print test tickets for all retrieved students, select **All Test Tickets**.

4. In the new browser window that opens displaying a layout for selecting the printed layout (see Figure 25), verify **Test Tickets** is selected in the *Print Options* section.

5. Select the layout you require, and then select **Print**.

Your browser downloads the generated PDF.

Figure 25. Layout Model for Test Tickets



How district-level users print test tickets from roster lists

1. From the **Print Test Tickets** task menu on the TIDE dashboard, select **Print from Roster List**. The ***View/Edit Rosters*** page appears.

2. Retrieve the rosters for which you want to print test tickets by filling out the search criteria and selecting **Search**.

3. Select the column headings to sort the retrieved rosters in the order you want the test tickets printed.

4. Do one of the following:

   - Mark the checkboxes for the rosters you want to print.
   - Mark the checkbox at the top of the table to print tickets for all retrieved rosters.

   When printing multiple class groups, the total number of students included in the rosters should not exceed 1000.

5. Select and then select **Test Tickets**. A layout model appears for selecting the printed layout (see Figure 25).

6. Verify **Test Tickets** is selected in the *Print Options* section.

7. Select the layout you require, and then select **Print**.

Your browser downloads the generated PDF.

How District-level Users Manage Appeal Requests

In the normal flow of a test opportunity, a student takes the test in TDS and then submits it. Next, TDS forwards the test for scoring, and then the test scores are reported in Reporting.

Appeal requests are a way of interrupting this normal flow. A test administrator-level user may want to request to re-open a test because of a hardware malfunction or a timeout issue.

For a full list of appeal request types, see the table "List of Appeal Request Types" in the appendix.

For a list of appeal request statuses, see the table "List of Appeal Request Statuses" in the appendix.

For a list of available appeal requests by test status, see the table "List of Appeal Requests by Test Status" in the appendix.

How district-level users add new appeal requests one at a time

You can create an appeal request for a given test result.

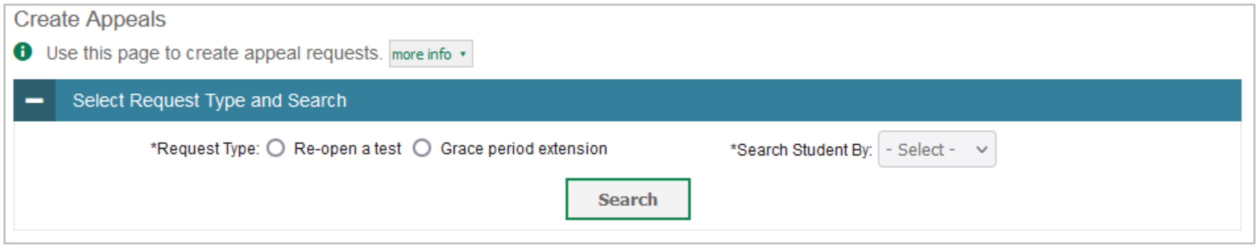
1. Retrieve the result for which you want to create an appeal request by doing the following:

   a. From the **Appeal Requests** task menu on the TIDE dashboard, select **Create Appeal Requests**. The ***Create Appeal Requests*** page appears (see Figure 26).

   b. Select a request type.

   c. From the drop-down lists and in the text field, enter search criteria.

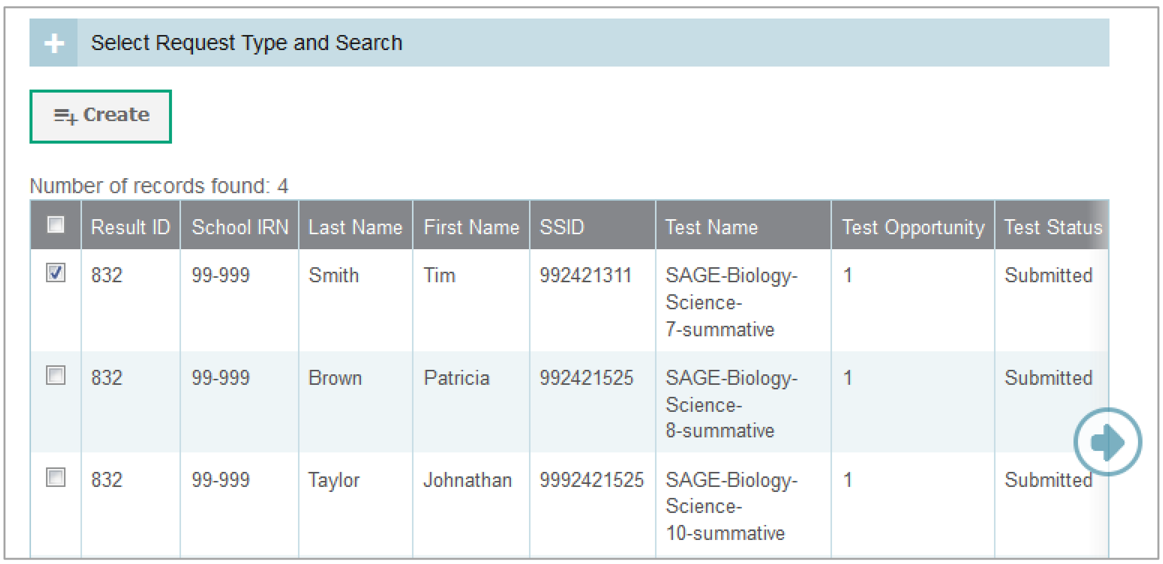
Figure 26. Create Appeal Requests Page



d. Select **Search**. TIDE displays the found results at the bottom of the ***Create Appeal Requests*** page (see Figure 27).

2. Mark the checkbox for each result for which you want to create an appeal, and then select **Create**.

Figure 27. Retrieved Test Results



	Result ID	School IRN	Last Name	First Name	SSID	Test Name	Test Opportunity	Test Status
☑	832	99-999	Smith	Tim	992421311	SAGE-Biology-Science-7-summative	1	Submitted
☐	832	99-999	Brown	Patricia	992421525	SAGE-Biology-Science-8-summative	1	Submitted
☐	832	99-999	Taylor	Johnathan	9992421525	SAGE-Biology-Science-10-summative	1	Submitted

3. Enter a reason for the request in the window that pops up.

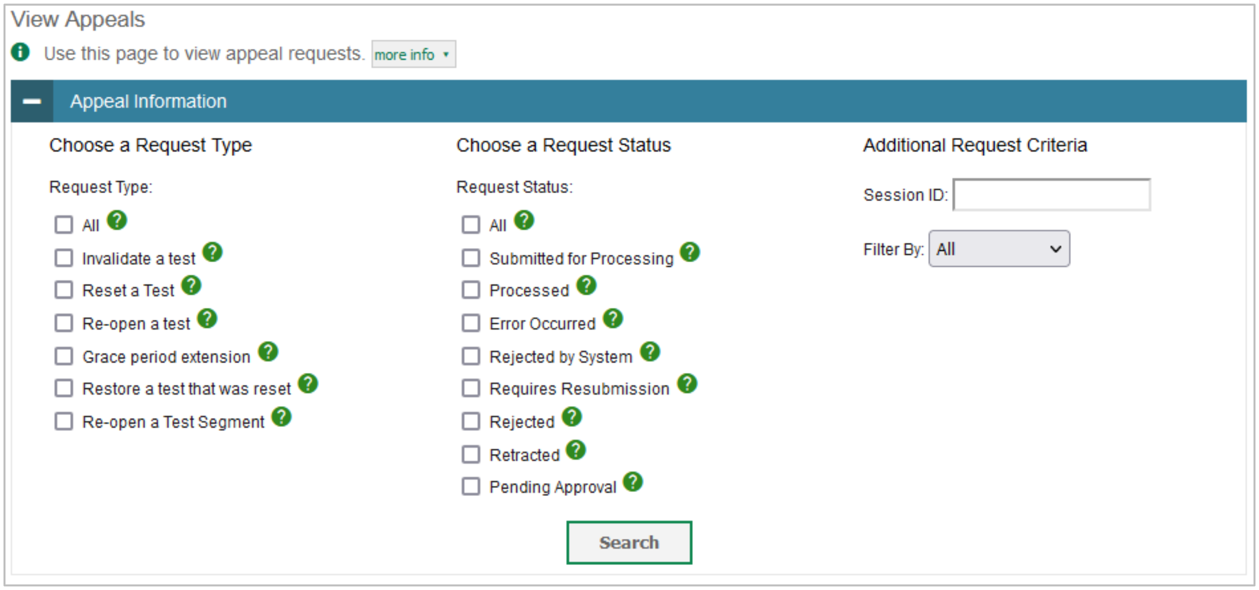
4. Select **Submit**. TIDE displays a confirmation message.

How district-level users view existing appeal requests one a at time

You can view and export existing appeal requests.

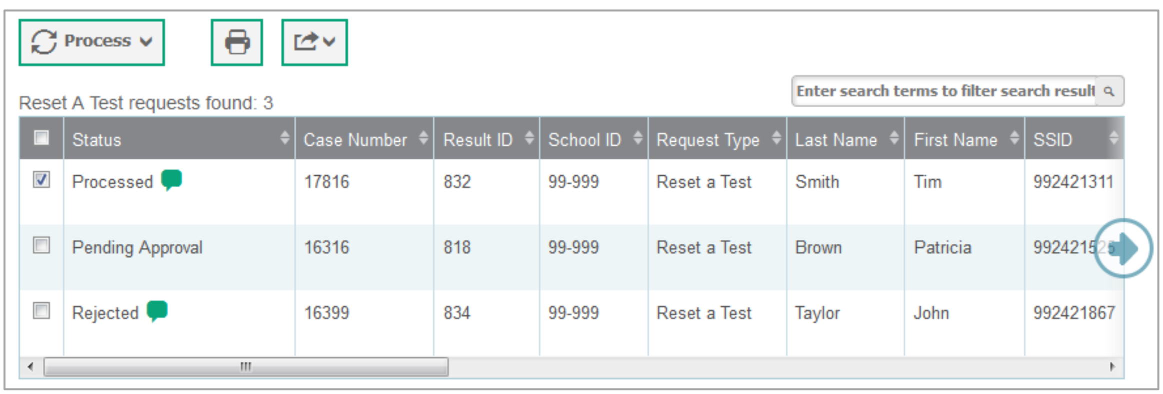
1. From the **Appeal Requests** task menu on the TIDE dashboard, select **View Appeal Requests**. The ***View Appeal Requests*** page appears (see Figure 28).

Figure 28. View Appeal Requests Page



2. Retrieve the appeal requests you want to view by filling out the search criteria and selecting **Search**. Figure 29 shows retrieved appeal requests.

Figure 29. Retrieved Appeal Requests



3. *Optional:* Review the initiator's reason for the appeal request by selecting in the Status column.

How district-level users retract appeal requests

You can retract appeal requests you created.

You cannot delete approved or rejected appeal requests. To delete such appeal requests, contact the Helpdesk.

1. From the **Appeal Requests** task menu on the TIDE dashboard, select **Approve Appeal Requests**. The ***Approve Appeal Requests*** page appears.

2. Retrieve the appeal requests you want to process by filling out the search criteria and selecting **Search**.

3. Mark the checkboxes for the requests you want to process.
4. Select **Process** above the table and select **Retract.**
5. Enter a reason for the requested action in the window that pops up.
6. Select **Submit**. TIDE displays a confirmation message.

TIDE removes the selected appeal requests from the list of retrieved requests.

How district-level users add or modify multiple appeal requests all at once

If you have many appeal requests to create, it may be easier to perform those transactions through file uploads. This task requires familiarity with composing comma-separated value (CSV) files or working with Microsoft Excel.

1. From the **Appeal Requests** task menu on the TIDE dashboard, select **Upload Appeal Requests**. The ***Upload Appeal Requests*** page appears.
2. Following the instructions in the section "How to add or modify multiple records at once" in the Introduction and using the table "Columns in the Appeal Request Upload File" in the appendix as a reference, fill out the Appeal Request template and upload it to TIDE.

How District-level Users Monitor Test Progress

The tasks available in the **Monitoring Test Progress** task menu allow you to generate various reports that provide information about a test administration's progress.

The following reports are available in TIDE:

- Plan and Manage Testing Report: Details a student's test opportunities and the status of those test opportunities. You can generate this report from the ***Plan and Manage Testing*** page.
- Test Completion Rates Report: Summarizes the number and percentage of students who have started or completed a test.
- Test Status Code Report: Displays all the test statuses for a test administration.

How district-level users view report of students' current test status

TIDE includes a Plan and Manage Testing report that details all of a student's test opportunities and the status of those test opportunities.

Because the report lists testing opportunities, a student can appear more than once on the report.

1. From the **Monitoring Test Progress** task menu on the TIDE dashboard, select **Plan and Manage Testing**. The ***Plan and Manage Testing*** page appears (see Figure 30).
2. In the *Search Students* panel, select the parameters for whose information to include in your report:
    a. From the **District** drop-down list, select a district if applicable.

b. From the **School** drop-down list, select a school if applicable. You may select one or more schools from this list. You may also select all schools. If you select more than 20 schools or if you select all schools and the district contains more than 20 schools, the report will only be available to export to the inbox.

c. *Optional*: If a single school was selected, choose a teacher from the **Teacher** drop-down list.

The **Teacher** drop-down list includes all school-level users, such as Teachers/Test Administrators (TA Role), Proctors (PR Role), and principal associated with the selected school. When you select a person from the **Teacher** drop-down list, TIDE performs a check to see if the person is associated with any roster. If no rosters exist for the selected person, no data is displayed when you generate the report. If the selected person has an associated roster, the plan and manage testing reports shows the test attempts of the students included in the roster.

If you do not select any person from the **Teacher** drop-down list and use the default value of **All** to generate the report, you will see all the tests taken in that school, irrespective of roster associations.

It is important to note that the TA Name displayed on the Plan and Manage Testing report does not imply the name of the teacher. The TA is the person who conducts the test. This can be the same as the teacher or it can mean a different person.

d. *Optional:* In the *Student's Last Name* field, enter a student's last name.

e. *Optional:* In the *Student's First Name* field, enter a student's first name.

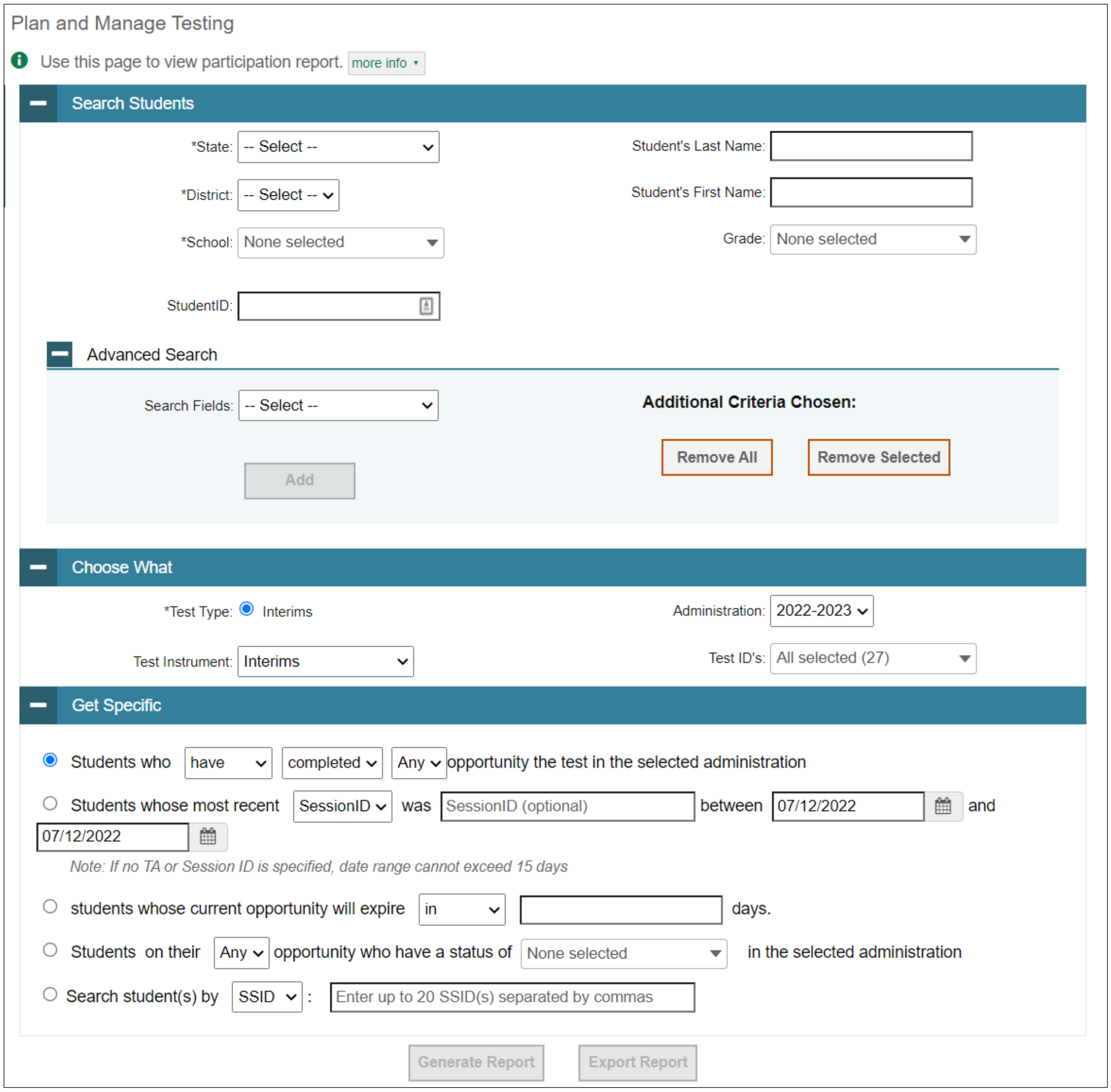
f. *Optional:* In the *SSID* field, enter a SSID.

g. *Optional:* From the **Grade** drop-down list, select a grade. You may select one, multiple, or all grades from this list.

3. In the *Choose What* panel, select the parameters for which tests to include in your report:

a. From the **Test** drop-down list, select a test category.

Figure 30. Plan and Manage Testing Page



b. From the **Test Instrument** drop-down list, select an assessment.

c. *Optional*: From the **Test Names** drop-down list, select the test for which you want to generate the report. You may select one, multiple, or all from this list.

d. *Optional*: From the **Filter By** drop-down list, select a specific test accommodation or demographic to filter the report.

- If you select a test accommodation or demographic, a *Values* field is displayed. Select the required filter criteria from the available options.

4. In the *Get Specific* panel, select the radio button for one of the options and then set the parameters for that option. The following options are available (parameters for each option are listed in {brackets}):

   - Students who {have/have not} {completed/started} the {1st/2nd/Any} opportunity in the selected administration.
     - Search for students who have not started the 1st/Any opportunity will return results for students who have not started their first opportunity on the selected test.
   - Students whose current opportunity will expire {in/between} {number/range} days.
     - If you select "in", you may enter any number in the displayed text box to determine tests expiring in the specified number of days. You may also enter 0 to see opportunities that expire that day.
     - If you select "between", you may enter two numbers in the displayed text boxes to signify a range of days (such as 1-3).
   - Students on their {1st/2nd/Any} opportunity in the selected administration, and have a status of {student test status}.
   - Students who have a status of {student test status} in the selected administration.
   - Students whose most recent {Session ID/TA Name} was {Optional Session ID/TA Name} between {start date} and {end date}.
   - Search student(s) by {SSID/Name}: {SSID/Student Name}

5. Do one of the following:

   - To view the report on the page, select Generate Report. If you are viewing a report for more than 20 schools, use this option and then select **Export to Inbox**.
   - To open the report in Microsoft Excel, select Export Report. If you are viewing a report for more than 20 schools, this option will be disabled.

Figure 31. Plan and Manage Testing Report

Number of records found: 2

Name	SSID	Enrolled Grade	Restricted Subjects	Current LEP	Test	Language
Smith, Ben	9999992563	03	ELA	N	Grade 3 Science	ENU
Garcia, Matt	9999992311	03	Social Sciences	Y	Grade 3 Mathematics	ENU

For descriptions of the columns in this report, see the table "Columns in the Plan and Manage Testing Report" in the appendix.

How district-level users view report of test completion rates

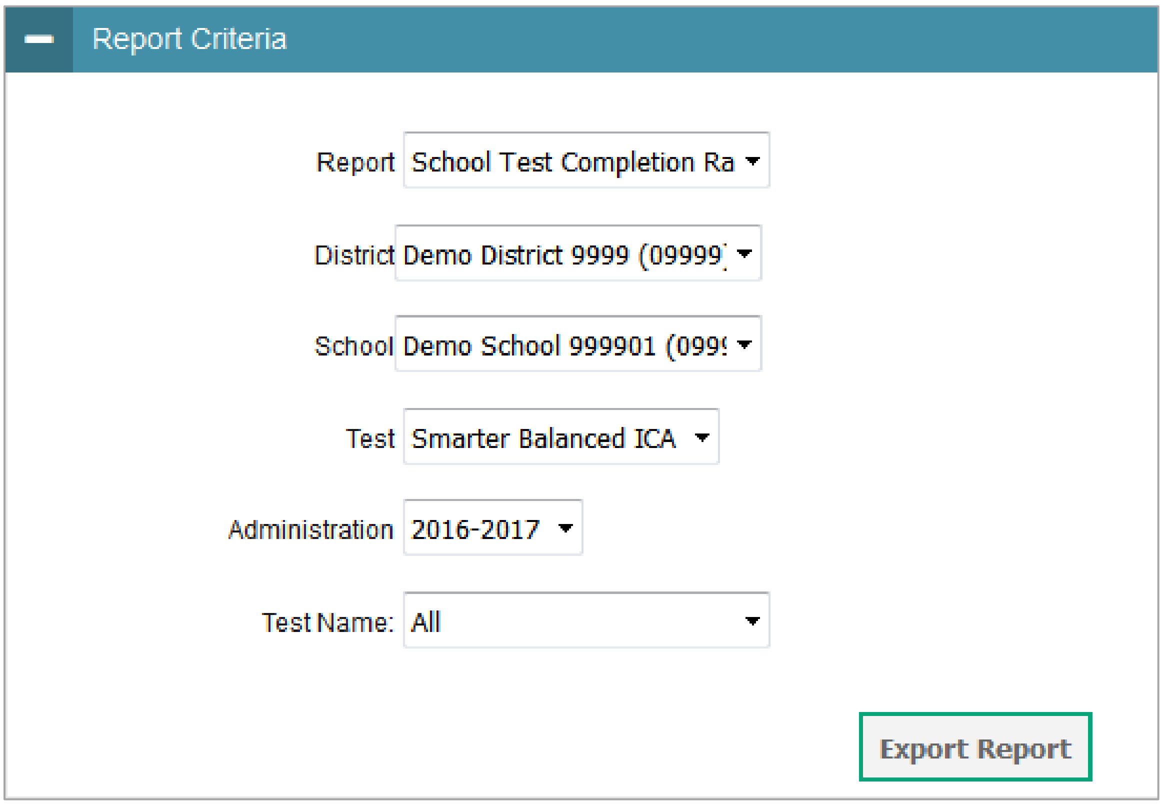
The Test Completion Rate report summarizes the number and percentage of students who have started or completed a test.

1. From the **Monitoring Test Progress** task menu on the TIDE dashboard, select **Test Completion Rates**. The ***Test Completion Rates*** page appears.

2. In the *Report Criteria* panel (see Figure 32), select the parameters for which tests to include in your report.

3. *Optional*: To disaggregate the report by student's enrolled grade, mark the *Show by Grade* checkbox.

Figure 32. Test Completion Rates Page: Report Criteria



4. To open the report in Microsoft Excel, select **Export Report**. Figure 33 displays a sample Test Completion Rate report.

5. For a description of the columns in this report, see the table "Columns in the Test Completion Rate Report" in the appendix.

Figure 33. Test Completion Rate Report

Number of records found: 2

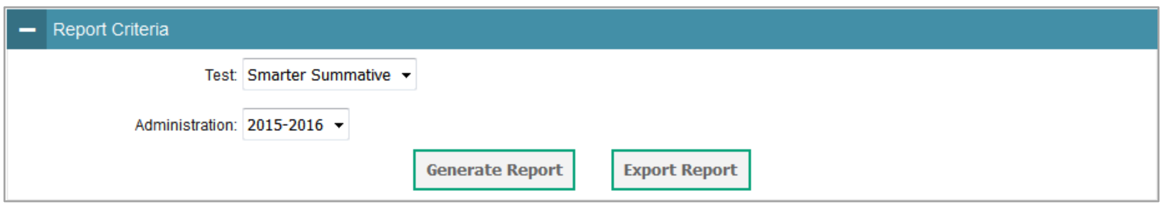
Date	Test Name	Opportunity	Total Student	Total Student Started	Total Student Completed	Percent Started	Percent Completed
02/08/2016	Grade 1 ELPA21 All Domains	1	7842	0	0	0.00%	0.00%
02/08/2016	Grade 1 ELPA21 Listening	03	31	0	0	0.00%	0.00%

How district-level users view report of test status codes

The Test Status Code report displays all current test statuses for a test administration.

1. From the **Monitoring Test Progress** task menu on the TIDE dashboard, select **Test Status Code Report**. The ***Test Status Code Report*** page appears.
2. In the *Report Criteria* panel (see Figure 34), select search criteria for the test and administration.

Figure 34. Report Criteria Panel



3. Do one of the following:
   - To view the report on the page, select Generate Report.
   - To open the report in Microsoft Excel, select Export Report.

TIDE displays the tests and associated statuses (see Figure 35).

Figure 35. Test Status Code Report

Student Name	SSID	Test Name	Test Status	Opportunity	Date Started	Assigned School ID	Assigned School Name
LastName, Fname		Grade 3 ELA - Interim Assessment (ICA)		0		9999_9999	Demo Institution 9999_9999
LastName, Fname		Grade 3 ELA - Interim Assessment (ICA)		0		9999_9999	Demo Institution 9999_9999

For a description of the columns in this report, see the table "Columns in the Test Status Code Report" in the appendix.

For a description of each status that a test opportunity can have, see the table "Test Opportunity Status Descriptions" in the appendix.

How School-level Users Perform Tasks in TIDE

School-level users (e.g., Building Coordinators (BC) and Proctors (PR Role) (TA)) have access to many of the same tasks as district-level users and perform these tasks the same way a district-level user performs them. For these tasks, this section of the guide refers school-level users back to the instructions presented in the district-level user section.

How School-level Users Perform Tasks in TIDE Before Testing Begins

Before testing begins, school-level users must perform the following tasks in TIDE:

- Set up **user accounts** for Teachers/Test Administrators (TA Role) and Proctors (PR Role) so they can sign in to TIDE and other CAI systems. If Teachers/Test Administrators (TA Role) or Proctors (PR Role) do not have accounts set up in TIDE, they will not be able to access any CAI systems or administer tests.
- Update **student test settings** to review test settings and ensure correct tests are assigned. If student accounts are not set up in TIDE in the correct test administration before testing begins, those students will not be able to test.
- Set up **rosters** so Reporting can display scores at the classroom, school, district, and state levels.

How School-level Users Set up User Accounts in TIDE

School-level users must set up user accounts in TIDE for Teachers/Test Administrators (TA Role) and Proctors (PR Role)proctors. If Teachers/Test Administrators (TA Role) and Proctors (PR Role)proctors do not have user accounts set up in TIDE before testing begins, they will not have access to any CAI systems or be able to administer tests.

Like district-level users, school-level users can add or modify user accounts one at a time or multiple user accounts all at once through file upload. These tasks can be performed following the procedure as described in the section "How District-level Users Set Up User Accounts in TIDE." For detailed information, please refer to the following sections:

- How district-level users add new user accounts one at a time
- How district-level users modify existing user accounts one at a time
- How district-level users add or modify multiple user accounts all at once

How School-level Users Register Students for Testing

School-level users can edit student test accessibility features (e.g., universal tools, designated supports, and accommodations) and test settings needs for testing. If students do not have the correct test settings assigned, they may not be able to sign in to a test.

All test settings must be enabled in TIDE prior to a student entering the test in order for them to be available. Some test settings cannot be edited by a user with the TA or PR-level role. You must have a AR/STC/BC role to edit test settings in TIDE. ⓘ***NOTE*** All accommodations require an IDEA/IEP or Section 504 plan. This flag must be set for a student before an accommodation can be updated in TIDE. Click here to review the Smarter Balanced Usability, Accessibility, and Accommodations Guide (UAAG), MSA UAAG, or Alternate MSA UAAG.

Turn on Student Test Settings at the Local Level

Accommodations are NOT transferred from the state AIM/Infinite Campus into the Test Information and Distribution Engine (TIDE). Per student privacy policies, a student is entered into TIDE as either having or not having an IEP/504. If a student has an IEP/504 marked in TIDE this enables the accommodations to be turned on. If the local district does not turn on individual student accommodations, then the student will not have accommodations available on the assessment(s).

Needs-Based Approach Towards Accessibility

Schools must administer statewide assessments with or without accommodations based on individual student needs consistent with all state and federal laws and regulations (ARM 10.56.104). Each school must support their students with accessibility tools per the student's educational plan. The supports that are provided on state assessments must be familiar to the student and match those supports and accommodations that are provided for classroom instruction and assessments throughout the school year. For additional guidance, please refer to Montana's Three Tiers of Accessibility.

These tasks can be performed following the procedure as described in the section "How District-level Users Update Students for Testing." For detailed information, please refer to the following sections:

- How district-level users specify student accommodations and test tools
- How district-level users upload student accommodations and test tools

How School-level Users Manage Rosters

School-level users can manage rosters for students in their school. These rosters are then sent to Reporting so those systems can display scores.

Like district-level users, school level users can add or modify rosters one at a time or all at once through file upload. These tasks can be performed following the procedure in the section "How District-level Users Manage Rosters." For detailed information, please refer to the following sections:

- How district-level users add new rosters one at a time
- How district-level users modify existing rosters one at a time
- How district-level users add or modify multiple rosters all at once

How School-level Users Use TIDE During Test Administration

During testing, school-level users can perform the following tasks in TIDE:

- Print **test tickets** to help students log in to tests.
- Add, modify, and upload appeal requests.
- View reports of students' current test statuses, test completion rates, and test status codes.

How School-level Users Print Test Tickets

School-level users can print test tickets for students in their school. Test tickets are hard-copy forms that includes a student's username for logging in to a test.

Test tickets can be printed by following the procedure in the section "How District-level Users Print Test Tickets." For detailed information, please refer to the following sections:

- How district-level users print test tickets from student lists
- How district-level users print test tickets from roster lists

How School-level Users Manage Appeal Requests

School-level users can manage appeal requests for students in their school.

Like district-level users, school-level users can add or modify appeal requests one at a time or all at once through file upload. These tasks can be performed by following the procedure in the section "How District-level Users Manage Appeal Requests." For detailed information, please refer to the following sections:

- How district-level users add new appeal requests one at a time
- How district-level users view existing appeal requests one at a time
- How district-level users add or modify multiple appeal requests all at once

How School-level Users Monitor Test Progress

Like district-level users, school-level users can view reports of students' current test statuses, test completion rates, and test status codes. These tasks can be performed by following the procedure in the section "How District-level Users Monitor Test Progress." For detailed information, please refer to the following sections:

- How district-level users view report of students' current test status
- How district-level users view report of test completion rates
- How district-level users view report of test status codes

How Teachers/Test Administrators (TA Role) and Proctors (PR Role) Proctors Perform Tasks in TIDE

Teachers/Test Administrators (TA Role) and Proctors (PR Role) have access to some of the same tasks as district-level and school-level users and perform these tasks the same way a district-level or school-level user performs them. For these tasks, this section of the guide refers Teachers/Test Administrators (TA Role) and Proctors (PR Role) back to the instructions presented in the district-level user section.

How Teachers/Test Administrators (TA Role) and Proctors (PR Role) Perform Tasks in TIDE Before Testing Begins

Before testing begins, Teachers/Test Administrators (TA Role) and Proctors (PR Role) proctors can perform the following tasks in TIDE:

- View **user accounts** to verify their own account information.
- View **student test settings** to ensure student details are properly entered into TIDE and edit student Universal Tools and Designated Supports as outlined in the UAAG, if necessary. All TA and PR users will need to contact their school-level user to turn on/edit student any accommodations. If student accounts are not set up in TIDE in the correct test administration before testing begins, those students will not be able to test.
- Set up **rosters** so Reporting can display scores at the classroom, school, district, and state levels.

How Teachers/Test Administrators (TA Role) and Proctors (PR Role) View User Accounts in TIDE

Teachers/Test Administrators (TA Role) and Proctors (PR Role) can view their own user account information in TIDE by selecting **Manage Accounts** from the banner.

How Teachers/Test Administrators (TA Role) and Proctors (PR Role) Manage Student Information

Teachers/Test Administrators (TA Role) and Proctors (PR Role) can view student accounts and student distribution reports by selecting the **Student** task menu, selecting **View Students**, filling out the search criteria, and selecting **Search**. Search results can be viewed in TIDE or exported to the inbox.

Like district- and school-level users, Teachers/Test Administrators (TA Role) and Proctors (PR Role) can also specify students' accommodations and test tools by following the procedure in the section "How district-level users specify student accommodations and test settings."

How Teachers/Test Administrators (TA Role) and Proctors (PR Role) Manage Rosters

Teachers/Test Administrators (TA Role) and Proctors (PR Role) can manage rosters for students in their school. These rosters are then sent to Reporting so those systems can display scores.

Like district- and school-level users, Teachers/Test Administrators (TA Role) and Proctors (PR Role) can add or modify rosters one at a time or all at once through file upload. These tasks can be performed following the procedure in the section "How District-level Users Manage Rosters."

How Teachers/Test Administrators (TA Role) and Proctors (PR Role) Use TIDE During Testing

During testing, Teachers/Test Administrators (TA Role) and Proctors (PR Role) can perform the following tasks in TIDE:

- Print **test tickets** to help students log in to tests.
- View reports of students' current test statuses, test completion rates, and test status codes.

How Teachers/Test Administrators (TA Role) and Proctors (PR Role) Print Test Tickets

Teachers/Test Administrators (TA Role) and Proctors (PR Role) can print test tickets for their students. Test tickets are hard-copy forms that includes a student's username for logging in to a test.

Test tickets can be printed by following the procedure in the section "How District-level Users Print Test Tickets." For detailed information, please refer to the following sections:

- How district-level users print test tickets from student lists
- How district-level users print test tickets from roster lists

How Teachers/Test Administrators (TA Role) and Proctors (PR Role) Monitor Test Progress

Like district- and school-level users, Teachers/Test Administrators (TA Role) and Proctors (PR Role) can view reports of students' current test statuses, test completion rates, and test status codes. These tasks can be performed by following the procedure in the section "How District-level Users Monitor Test Progress." For detailed information, please refer to the following sections:

- How district-level users view report of students' current test status
- How district-level users view report of students' current test status by student ID
- How district-level users view report of test completion rates
- How district-level users view report of test status codes

Appendix

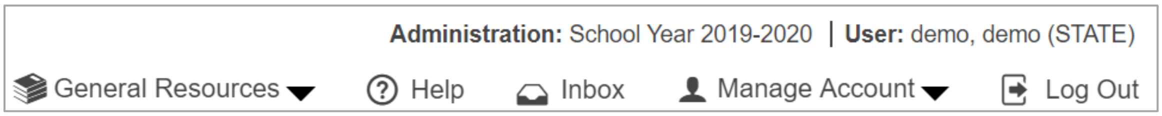
A

Accessibility Supports

TIDE provides a link to definitions for accessibility supports.

1. From the **General Resources** drop-down list in the banner (see Figure 36), select **Accessibility Supports**. The ***Accessibility Supports*** page appears.

Figure 36. TIDE Banner



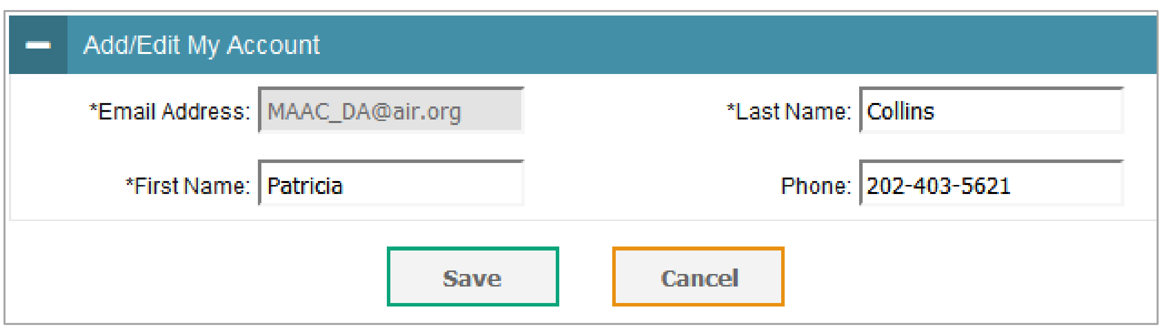
Account Information

You can modify your name, phone number, and other account information in TIDE. (To change your email address, your school or district assessment coordinator must create a new account with the updated email address.)

1. In the TIDE banner (see Figure 36), from the **Manage Account** drop-down list, select **My Contact**. The ***My Contact Information*** page appears (see Figure 37).
2. Enter updates as necessary.
3. Select **Save**.

TIDE saves your changes, and a confirmation message appears.

Figure 37. Fields in the My Contact Information Page



C

Changing Your Associated Test Administration, Institution, or Role

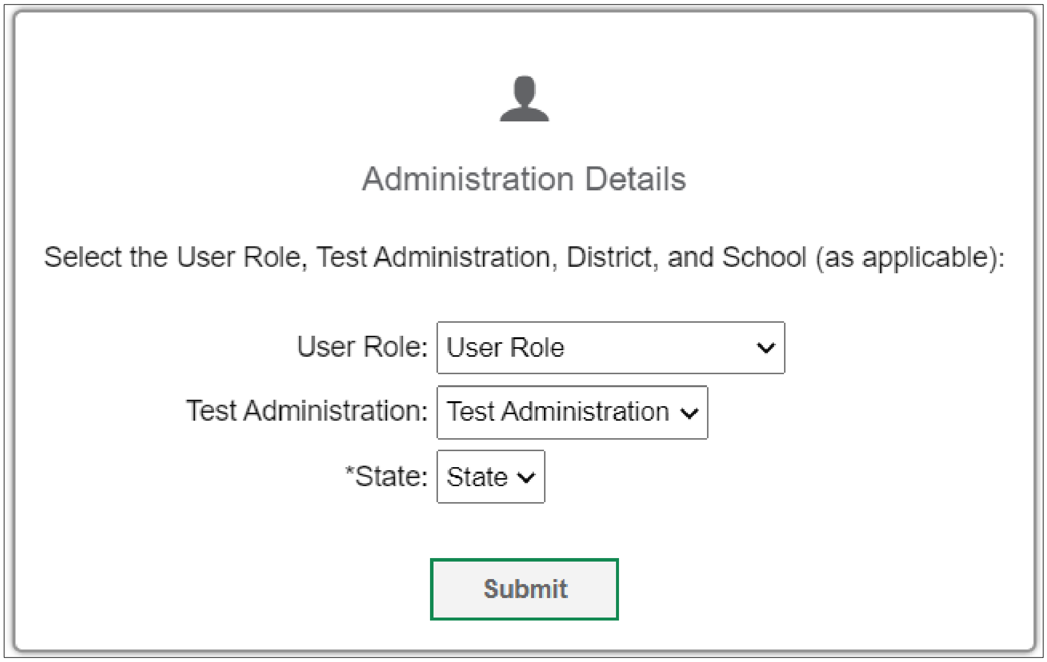
Depending on your permissions, you can switch to different schools, districts, and user roles in TIDE.

1. In the TIDE banner (see Figure 36), select **Change Role** from the **Manage Account** drop-down menu. The ***Administration Details*** window appears (see Figure 38).

1. Update the information as necessary.

2. Select **Submit**. A new home page appears that is associated with your selections.

Figure 38. Administration Details Window



Columns in the Appeal Request Upload File

You can use the information in the table below to add or modify multiple appeal requests all at once through file upload.

Column Name	Description	Valid Values

Type*	Type of appeal request.	One of the following:
		`Grace Period Extension`
		`Re-open a test`

Search Type*	Student field to search.	One of the following:
		`Result ID`
		`Session ID`
		`SSID`

Search Value*	Search value corresponding to the search type.	Up to 1,000 alphanumeric characters. The value must exist in TDS or TIDE. For example, specifying a result ID of 123456 requires that this result ID exist in TDS.

Column Name	Description	Valid Values
Reason*	Reason for creating appeal request.	Up to 1,000 alphanumeric characters.

*Required field.

Columns in the Interim Grades Upload File

You can use the information in the table below to upload interim grades.

Field Name	Description	Valid Values
SSID*	State-assigned student identifier.	Up to 10 numeric characters. Must be enrolled in your district.
Field*	Label used for the interim grade attribute.	Interim Testing Grade
Subject*	Subject of assessment.	One of the following: • Mathematics • English Language Arts • Science
Interim Grade*	Student's interim grade.	Any of the following: Grade level in the format "Grade 3, Grade 4, etc." None

*Required field.

Columns in the Plan and Manage Testing Report

You can use the information in the table below to view report of students' current test status through the Plan and Manage Testing module or when searching by student ID.

Attribute	Description
Name	Student's legal name (Last Name, First Name).
SSID	Student's Statewide Student Identifier number.
District Name	Name of the district associated with the record.
School Name	Name of the school associated with the record.
Enrolled Grade	The grade in which a student is enrolled.
Interim Test Grade	Indicates the interim grades set up for the student.
Test	Test name for this student record.
Language	The language setting that was assigned to the student (English or Spanish).

Attribute	Description
Results ID	The unique identifier linked to the student's results for that specific opportunity.
Opportunity	The opportunity number for that student's specific record.
Date Started	The date when the first test item was presented to the student for that opportunity.
Date Completed	The date when the student submitted the test for scoring.
TA Name	The test administrator-level user who created the session in which the student is currently testing (or in which the student completed the test).
Session ID	The Session ID to which the test is linked.
Status	The status for that specific opportunity. This column also shows the number of items to which a student has responded and the total number of items on the test.
Restarts	The total number of times a student has resumed an opportunity (e.g., if a test has been paused three times and the student has resumed the opportunity after each pause, this column will show three restarts). (This includes Restarts Within Grace Period—see below.)
Restarts Within Grace Period	The total number of times a student has resumed an opportunity within 20 minutes after a test was paused. For example, if a test has been paused three times and the student resumed the opportunity within 20 minutes of two pauses but 25 minutes after the third pause, this column shows two Restarts Within Grace Period). A student has a grace period of 20 minutes to pause the test at a test item and then resume the test at that same item. However, if a test is paused for more than 20 minutes, the test session will expire and the student will not be able to review any previous answers.
Last Activity	The date of the last activity for that opportunity or record. A completed test can still have activity as it goes through the QA and reporting process.
Expiration Date	The date the test opportunity expires.
Force Complete Date	The date a test expired and was force-completed.
Remote Session	Indicates if the test was taken in person or remotely.

Columns in the Roster Upload File

You can use the information in the table below to add or modify multiple rosters all at once.

Column Name	Description	Valid Values
District ID*	District associated with the roster.	District ID that exists in TIDE. Up to 20 characters.

Column Name	Description	Valid Values
School ID*	School associated with the roster.	School ID that exists in TIDE. Up to 20 characters. Must be associated with the district ID.
User Email ID*	Email address of the teacher associated with the roster.	Email address of a teacher existing in Reporting.
Roster Name*	Name of the roster.	Up to 20 characters.
SSID*	Student's unique identifier within the district.	Up to 30 alphanumeric characters.
Action	Action column to add or delete students from roster	Add – adds student to roster Delete – deletes student from roster

*Required field.

Columns in the Test Completion Rate Report

You can use the information in the table below to view report of test completion rates.

Column	Description
Date	Date and time that the file was generated.
Test Name	Grade, test, and subject that are being reported.
Test	Test that is being reported.
Administration	Administration that is being reported.
District Name	The name of the reported District.
District ID	The ID of the reported District.
School Name	The name of the reported school. This column is only included in the school-level report.
School ID	The ID of the reported school. This column is only included in the school-level report.
Opportunity	Test opportunity number that is being reported.
Total Student	Number of students with an active relationship to the school in TIDE.
Total Student Started	Number of students who have started the test.
Total Student Completed	Number of students who have finished the test and submitted it for scoring.
Percent Started	Percentage of students who have started the test out of the total number of students with an active relation to the school in TIDE.

Column	Description
Percent Completed	Percentage of students who have completed the test out of the total number of students with an active relation to the school in TIDE.
Test Family	Group of tests containing the test that is being reported.
Total Student Remote Session	Number of students who took the test remotely.
Percent Remote Session	Percentage of students who took the test remotely.
Enrolled Grade	Student's enrolled grade. This column is only populated if you marked the Show by Grade checkbox.

Columns in the Test Settings Upload File

You can use the information in the table below to upload student accommodations and test tools.

Column	Description	Valid Values
SSID*	Student's statewide identification number.	Ten digits.
Subject	Subject for which the tool or accommodation applies.	One of the following: ELA ELA-PT (Interim ONLY) Mathematics Science Alt Science
Tool Name	Name of the tool or accommodation.	See the table "Valid Values for Tool Names."
Value	Indicates if the tool or accommodation is allowed or disallowed, or the accommodation's appearance.	See the table "Valid Values for Tool Names."

*Required field.

Columns in the Test Status Code Report

You can use the information in the table below to view reports of test status codes.

Column	Description
Student Name	Student's name.
SSID	Student's Statewide Student Identifier number.

Column	Description
OppNum	Test opportunity number.
Test Name	Test in which student did not participate.
Test Status	Test's most recent status.
Date Started	Date student started the test.
Result ID	Unique ID for the item result.
Session ID	Unique ID for the test session.
Test Expiration Date	Date the test expired.

Columns in the User Upload File

You can use the information in the table below to add or modify multiple user accounts all at once through file upload.

Column	Description	Valid Values
DISTRICTID*	District associated with the user.	District ID that exists in TIDE, and must be associated with the user uploading the file. Up to 20 characters.
SCHOOLID	School associated with the user.	School ID that exists in TIDE, and must be associated with the user uploading the file. Up to 20 characters. Must be associated with the district ID. Can be blank when adding district-level users.
First Name*	User's first name.	Up to 35 characters.
Last Name*	User's last name.	Up to 35 characters.
Email Address*	User's email address.	Any standard email address. Up to 128 characters that are valid for an email address. This is the user's username for logging in to TIDE.
Phone Number	User's phone number.	Phone number in xxx-xxx-xxxx format. Extensions allowed.

Column	Description	Valid Values
Role*	User’s role. For an explanation of user roles, see [User Role Permissions]().	One of the following: AR—Authorized Representative STC—School Test Coordinator BC—Building Coordinator TA—Test Administrator PR—Proctor Must be lower in the hierarchy than the user uploading the file.
Action*	Indicates if this is an add, modify, or delete transaction.	One of the following: Add—Add new user or edit existing user record. Delete—Remove existing user record.

*Required field.

D

Deleting Records from TIDE

You can delete existing records for user and rosters from TIDE. For users with multiple roles, individual roles can be deleted without deleting the entire user account.

1. Retrieve the records you want to delete by following the procedure in the section Searching for Records in TIDE.
2. Do one of the following:
   - Mark the checkboxes for the record you want to delete.
   - Mark the checkbox at the top of the table to delete all retrieved records.
3. Select , and in the affirmation dialog box select **OK**.

E

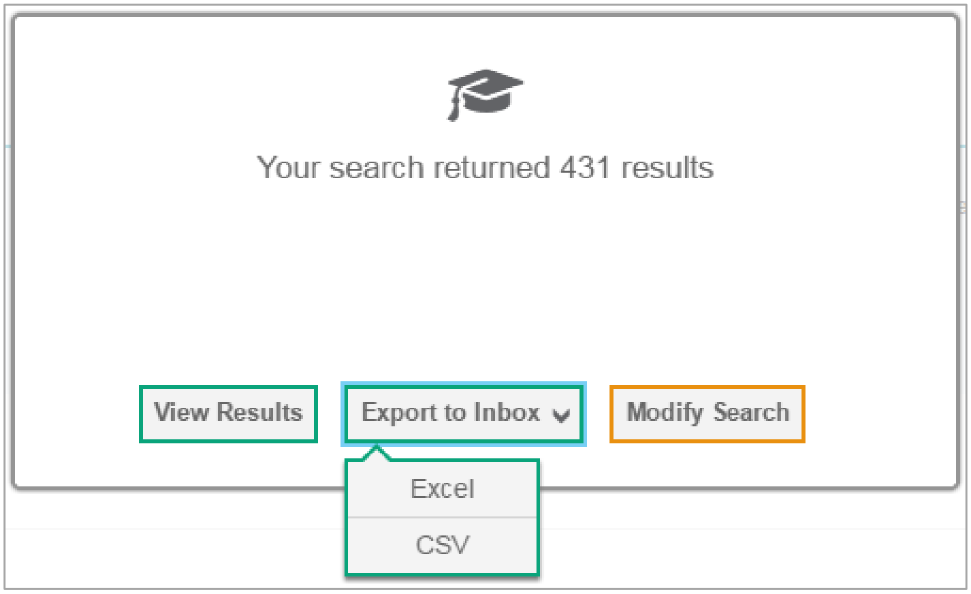
Exporting Records in TIDE

You can export search results for users, students, rosters, students’ test settings, test windows, and appeal requests to the inbox.

1. Retrieve the records you want to export by following the procedure in the section Searching for Records in TIDE.

2. In the search results pop-up window, select **Export to Inbox** and select the file format (CSV or Excel) in which the data should be exported. You can navigate away from the page and perform other tasks if required. When your file is available for download, you will receive an email to the email account registered in TIDE. After receiving the email, you can download the exported file from the Inbox.

Figure 39. Search Results



You can also export records from the search results grid.

1. Retrieve the records you want to delete by following the procedure in the section [Searching for Records in TIDE].

2. Do one of the following:

   - Mark the checkboxes for the record you want to export.
   - Mark the checkbox at the top of the table to export all retrieved records.

3. Select , and then select Excel or CSV.

F

Fields in the Demographics Panel

You can use the information in the table below to add new student accounts with permanent IDs one at a time or to modify existing student accounts one at a time.

Field	Description
Grade	Grade in which student is enrolled during the test administration.
Last Name*	Student's last name.

Field	Description
First Name*	Student's first name.
Middle Name	Initial of student's middle name.
SSID*	Student's Statewide Student Identifier (SSID) within the enrolled district.
Birth Date*	Student's date of birth.
Gender	Student's gender.
Section 504	Student's 504 status.
IDEA Indicator/IEP	Student's IDEA/IEP status.
LEP Status	Student's LEP status.
Economic Disadvantage Status	Student's economic disadvantage status.
Primary Disability Type	Student's primary disability.
Alternate Assessment	Student eligibility for the alternate assessment.
Race and Ethnicity	Student's race and ethnicity.

*Required field.

Fields in the Test Settings and Tools Panels

When you are "updating students for testing", you can use the information in the table below to enter the student's settings for each test.

Field	Description
Embedded Designated Supports	
Color Choices	List of available color settings.
Glossaries (English and Translated)	List of available languages for students to view a glossary for certain words in the test content.
Masking	Toggles the Masking tool on or off, allowing student to cover distracting regions of the test page.
Mouse Pointer [Size and Color]	List of available colors and sizes for mouse pointer.
Permissive Mode	Toggles Permissive Mode setting on or off, allowing student to use pre-approved hardware or software with secure browser.

Field	Description
Streamlined Mode	Toggles Streamlined Mode setting on or off, displaying test content vertically.
Text-to-Speech	Sets which test content is administered with the TTS accommodation.
Zoom	List of available zoom levels.
Embedded Accommodations	
Braille Transcript	Toggles Transcriptions tool on or off, displaying audio content transcripts.
Braille Type	List of available braille types for test content.
Closed Captioning (Listening Items)	Toggles the Closed Captioning tool on or off, displaying captions for audio content.
Language	List of available languages for test content.
Sign Language (Embedded)	Toggles the Sign Language tool on or off, displaying videos translating test content into ASL.
Speech-to-Text (Embedded)	Toggles the Speech-to-Text tool on or off, allowing student to dictate responses on constructed response items.
Non-Embedded Designated Supports	
Non-Embedded Designated Supports	List of available non-embedded designated supports.
Non-Embedded Accommodations	
Non-Embedded Accommodations	List of available non-embedded designated accommodations.
Print on Request	Toggles the Print on Request tool on or off, allowing students to request test content to be printed.

Fields in the View/Edit Users [User's Name] Page

You can use the information in the table below to modify existing user accounts.

Field	Description
Email Address*	Email address for logging in to TIDE.
Role*	User's role. For an explanation of user roles, see User Role Permissions.
District*	District associated with the user.
School*	School associated with the user.
First Name	User's first name.
Last Name	User's last name.
Phone	User's phone number.

Field	Description
TA Certified	Indicates if the user has been trained to use online assessment systems. Once the user completes the TA Certification Course this field will automatically populate with a flag.

*Required field.

H

Hand-Scoring Resources

TIDE provides resources you can use to prepare for scoring tests by hand.

1. From the **General Resources** drop-down list in the banner (see Figure 36), select **Interim Materials**. The ***Interim Materials*** page appears.

2. Select the download link for the required resource.

I

Inbox Files

When searching for users, students, students' test settings, test windows, and appeal requests, you can choose to export the search results to the Inbox. The shared Inbox serves as a secure repository that lists files containing the data that you have exported in TIDE and other CAI systems. When you choose to export search results to the Inbox, TIDE sends you an email when the export task is completed and the file is available in the Inbox for download.

The Inbox also lists any secure documents that have been externally uploaded to the Inbox and that you have privileges to view.

The files in the Inbox are listed in the order in which they were created. The file creation and file expiration dates appear, if applicable. The number of days remaining until a file expires is also displayed next to a file. By default, exported files are available for 30 days while secure documents are available for the period specified by the OPI. You can access the Inbox from any page in TIDE to either download the file or archive the file for future reference. You can also delete the files you have exported, provided you have not archived them.

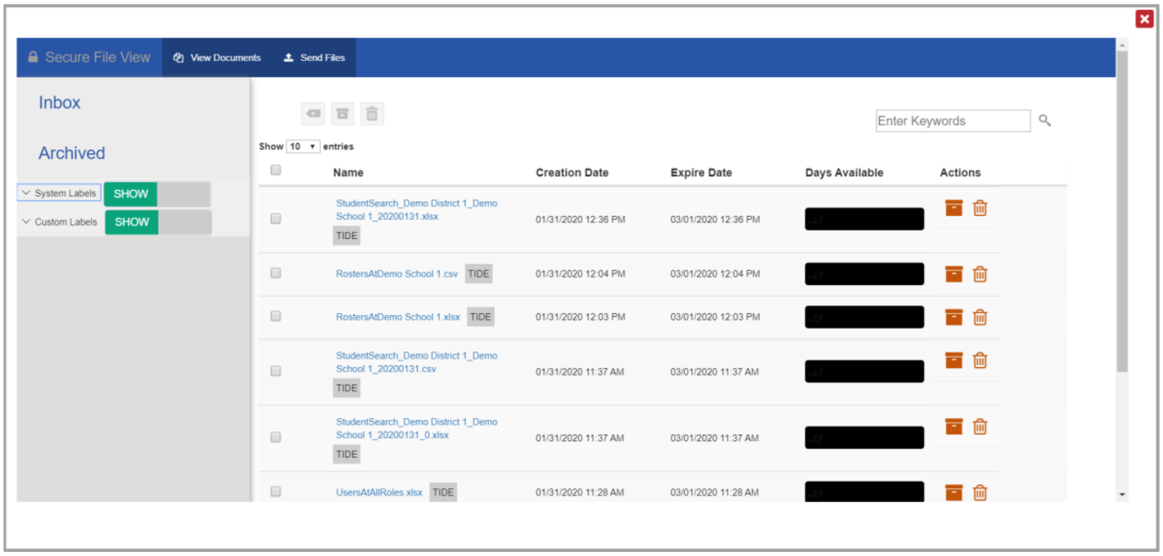
1. From the TIDE banner (see Figure 36), select **Inbox**. The ***Inbox*** page appears (see Figure 40). By default, TIDE displays the *View Documents* tab.

2. *Optional*: Select the file view from the available tabs:

   - **Inbox**: This is the default view and displays all the files except for the ones that you have archived.

   - **Archived**: Displays the files that you have archived.

Figure 40. Inbox Window: View Documents Tab



3. *Optional*: To filter the files by keyword, enter a search term in the text box above the list of files. TIDE displays only those files containing the entered file name.

4. *Optional*: To hide or display system labels, toggle 

.

5. *Optional*: To hide files with a system label, unmark the checkbox for that system label.

6. *Optional*: To hide or display custom labels, toggle 

.

7. *Optional*: To hide files with a custom label, unmark the checkbox for that custom label.

8. Do one of the following:

   - To download a file, select the file name.
   - To add a new custom label or apply an existing custom label, select .
     - To apply a new custom label, mark the checkbox, enter a new custom label in the text box, and select **Save New Label**.
     - To apply an existing custom label, mark the checkbox, enter an existing custom label in the text box, and select **Apply Label**.
   - To archive a file, select .
   - To delete a file, select .

About File Deletion

- Archived files cannot be deleted.

- You can delete files that you have exported, but you cannot delete secure documents uploaded to the Inbox by admin users.

L

List of Appeal Requests by Test Status

You can use the information in the table below to manage appeal requests. Appeal request can only be generated for a test if the test opportunity is in the indicated test status.

Test Status	Re-open a test	Grace Period Extension
Approved		
Completed	✓	
Denied		✓
Expired	✓	
Paused		✓
Pending		
Processing		
Reported	✓	
Review		
Scored	✓	
Started		
Submitted	✓	
Suspended		
Invalidated	✓	

List of Appeal Request Statuses

You can use the information in the table below to manage appeal requests. The table describes the status of an appeal request once submitted.

Appeal Request Status	Description of Status
Error Occurred	An error occurred while the appeal request was being processed.

Appeal Request Status	Description of Status
Pending Approval	Appeal request is pending approval.
Processed	Appeal request was successfully processed and the test opportunity has been updated.
Rejected	Another user rejected the appeal request.
Requires Resubmission	Appeal request must be resubmitted.
Retracted	Originator retracted the appeal request.
Submitted for Processing	Appeal request submitted to Test Delivery System for processing.

List of Appeal Request Types

You can use the information in the table below to manage appeal requests.

Appeal requests must be submitted at least one day prior to the end of a test window so that students can complete their test opportunity.

Type	Description
Re-open a test	Reopens a test that was completed, invalidated, or expired.
Grace Period Extension (GPE)	Allows the student to review previously answered questions upon resuming a test or test segment after expiration of the pause timer. For example, a student pauses a test, and a 20-minute pause timer starts running. The following scenarios are possible: • If resuming the test within 20 minutes, student can review previously answered questions. • Without a GPE, student resuming the test after 20 minutes cannot review previously answered questions—student can only work on unanswered questions. Upon receiving a GPE, student can review previously answered questions upon resuming the test. The normal pause rules apply to this opportunity.

P

Password Information

Your username is the email address associated with your account in TIDE. When you are added to TIDE, you receive an activation email containing a temporary link to the ***Reset Your Password*** page. To activate your account, you must set your password within 15 minutes of the email being sent.

- **If your first temporary link expired:**

In the activation email you received, select the second link provided and proceed to request a new temporary link.

- **If you forgot your password:**

On the ***Login*** page, select **Forgot Your Password?** and then enter your email address in the *Email Address* field. You will receive an email with a new temporary link to reset your password.

- **If you did not receive an email containing a temporary link or authentication code:**

Check your spam folder to make sure your email program did not categorize it as junk mail. If you still do not have an email, contact your School or District Test Coordinator to make sure you are listed in TIDE.

- **Additional help:**

If you are unable to log in, contact the Montana Assessment Helpdesk for assistance. You must provide your name and email address. Contact information is available in the User Support section of this user guide.

Printing Records in TIDE

1. Retrieve the records you want to print by following the procedure in the section Searching for Records in TIDE.

2. Do one of the following:

   - To print some records, mark the checkboxes for the records you want to print, select , select My Selected, and then select Print.

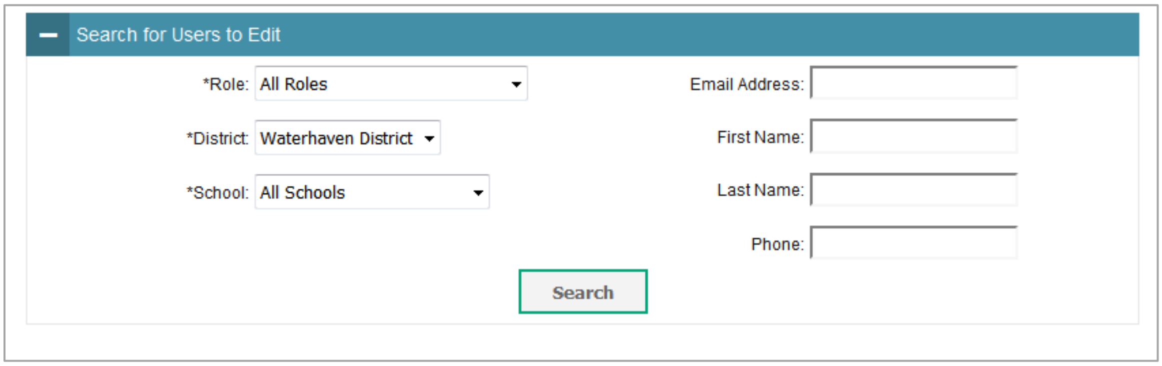
   - To print all records, select , select All, and then select Print.

S

Searching for Records in TIDE

Many tasks in TIDE require you to retrieve a record or group of records (for example, locating a set of users to work with when performing the **View/Edit/Export Users** task). For such tasks, a search panel appears when you first access the task page (see Figure 41). This section explains how to use this search panel and navigate search results.

Figure 41. Sample Search Panel



1. In the search panel, enter search terms and select values from the available search parameters, as required. Some fields may allow you to select multiple values. For example, the school drop-down lists on the student search pages will allow you to select one, multiple, or all values. Similarly, the **Test ID** drop-down list on the ***Plan and Manage Testing*** page will allow you to select one, multiple, or all values.

   The search parameters available in the search panel depend on the record type. Required search parameters are marked with an asterisk.

2. *Optional*: If the task page includes an additional search panel, select values to further refine the search results:

   - To include an additional search criterion in the search, select it and select Add or Add Selected as available

   - Optional: To delete an additional search criterion, select it and select Remove Selected. To delete all additional search criteria, select Remove All.

3. Select **Search**.

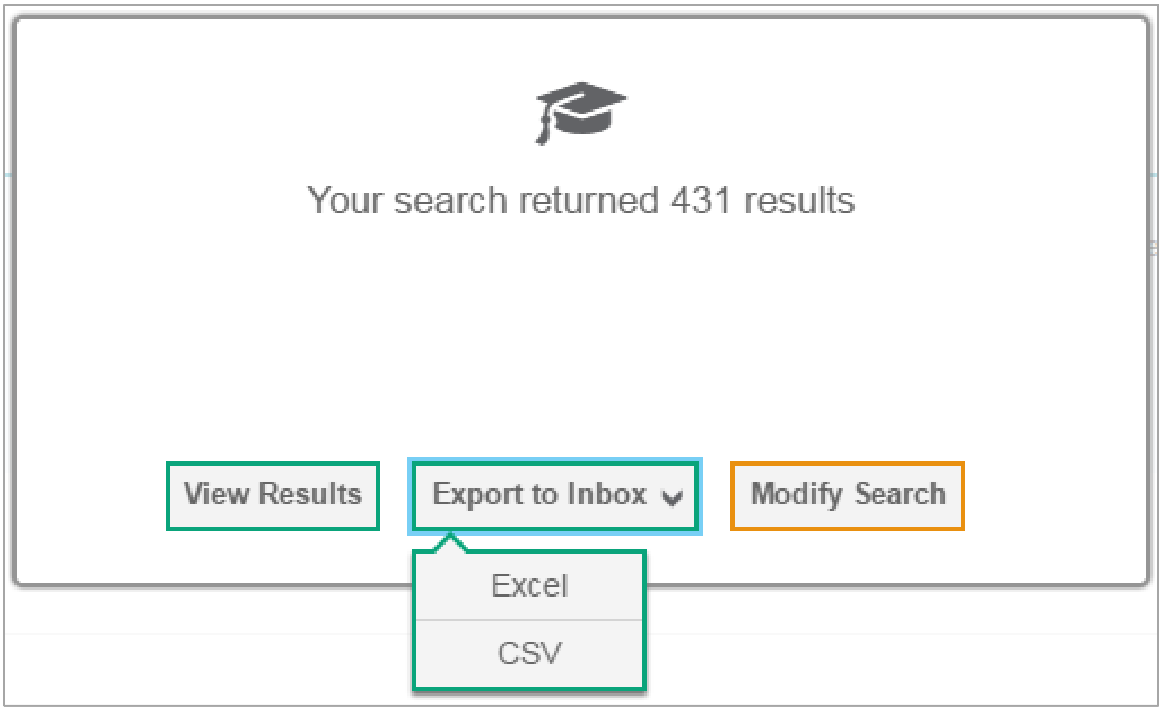
   - If searching for users, students, students' test settings, and appeal requests, proceed to the next step.

   - If searching for other types of records, such as rosters, skip to step 7.

4. In the search results pop-up window (see Figure 42) that indicates the number of records that matched your search criteria and provides you with options to view or export the records or modify your search parameters, do one of the following:

   - To view the retrieved records on the page, select **View Results**. Continue to step 7. This option is not available if TIDE detects that this action might adversely affect its performance.

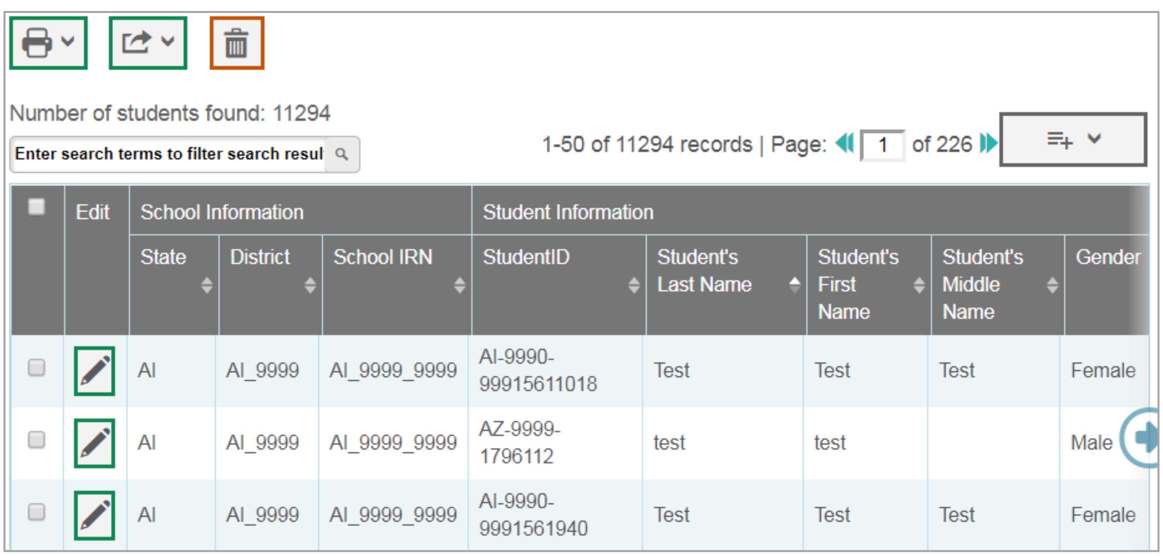
Figure 42. Search Results Pop-up Window



- To export the retrieved results to the Inbox, select **Export to Inbox** and select the file format (CSV or Excel) in which the data should be exported. You can navigate away from the page and perform other tasks if required. When your file is available for download, you will receive an email to the email account registered in TIDE. After receiving the email, you can download the exported file from the Inbox (see Inbox Files).
- To return to the page and modify your search criteria, select **Modify Search**. Repeat steps 1–3.

5. The list of retrieved records appears below the search panel (see Figure 43).

Figure 43. Sample Search Results



6. *Optional*: To filter the retrieved records by keyword, enter a search term in the text box above the search results and select . TIDE displays only those records containing the entered value.

7. *Optional*: To sort the search results by a given column, select its column header.

   - To sort the column in descending order, select the column header again.

8. *Optional*: If the table of retrieved records is too wide for your browser window, you can select and at the sides of the table to scroll left and right, respectively.

9. *Optional:* If the search results span more than one page, select or to view previous or next pages, respectively.

10. *Optional*: To hide columns, select (if available) and uncheck the checkboxes for the columns that you wish to hide. To show columns again, mark the applicable checkboxes.

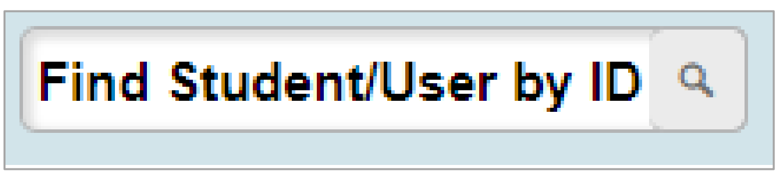
Searching for Students or Users by ID

A *Find Student/User by ID* field appears in the upper-right corner of every page in TIDE. You can use this field to navigate to the ***View and Edit Student*** or ***View/Edit User: [User's Name]*** form for a specified student or user.

1. In the *Find Student/User by ID* field, enter a student's SSID or a user's email address. The SSID or email address must be an exact match; TIDE does not search by partial SSID or email address.

2. Select . The ***View and Edit Student*** or ***View/Edit User: [User's Name]*** form for that student or user appears.

Figure 44. Find Student/User by ID Search Field



Sending Files from the Inbox

You can send a file or files from TIDE to individual recipients by email address or to groups of recipients by user role.

1. From the TIDE banner (see Figure 36), select **Inbox**. The ***Inbox*** page appears (see Figure 40). By default, TIDE displays the *View Documents* tab.

2. Select the **Send Files** tab. The ***Send Files*** page appears (see Figure 45).

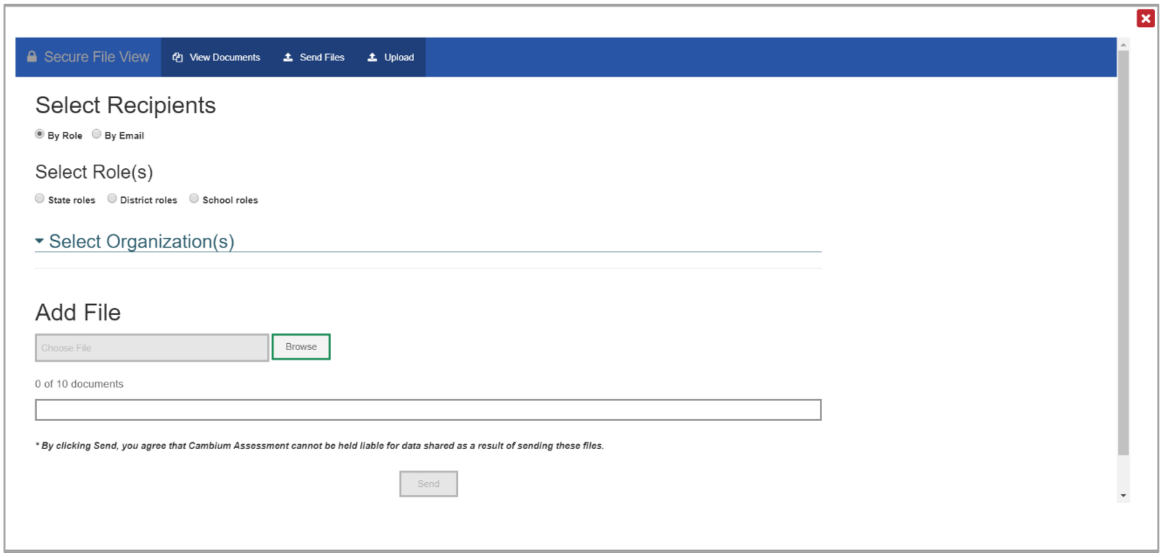
3. In the *Select Recipients* field, do one of the following:

   - Select **By Role** to send a file or files to a group of users by user role.

- Select **By Email** to send a file or files to a single recipient by email address.

If you select **By Email**, skip to step 7.

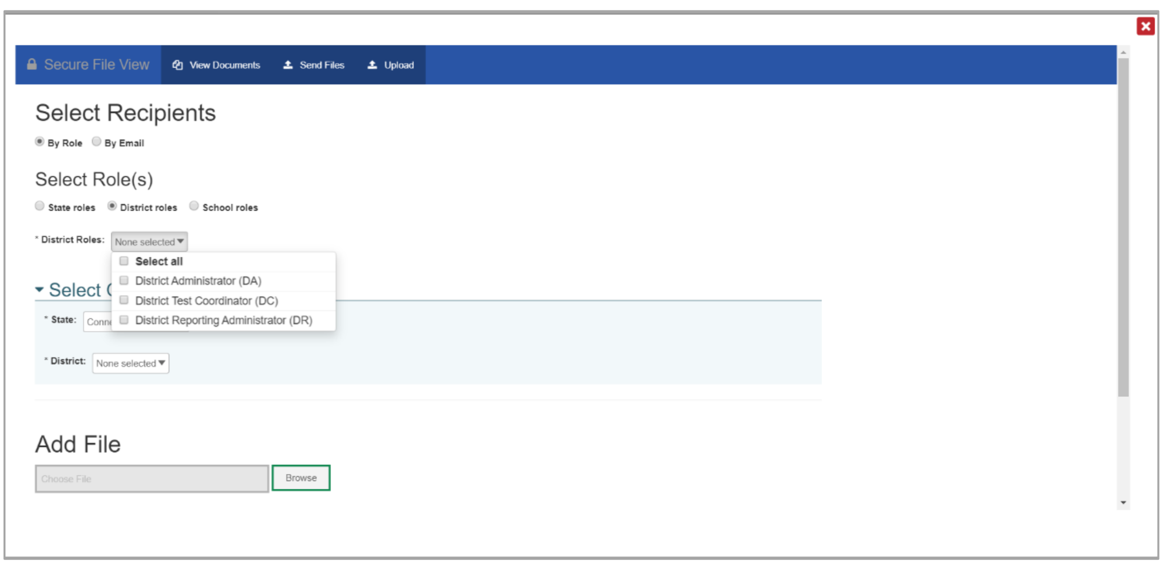
Figure 45. Inbox Window: Send Files Tab



4. In the *Select Role(s)* field, select the role group to which you want to send a file or files. A drop-down list appears (see Figure 46).

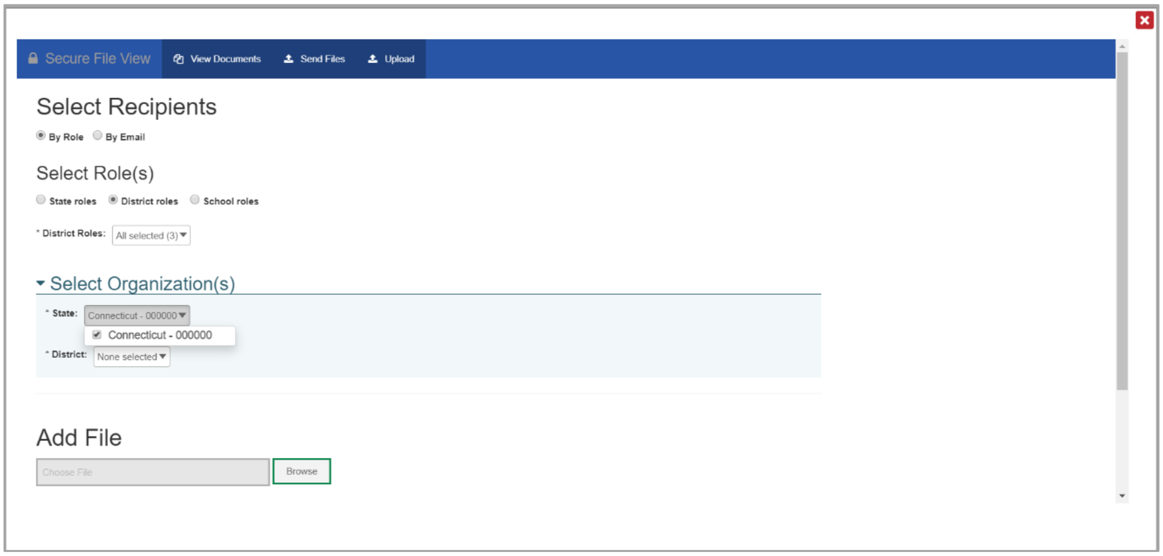
5. From the drop-down list, select the role(s) to which you want to send a file or files. You can choose **Select all** to send a file or files to all roles in the selected role group.

Figure 46. Inbox Window: Send Files Tab with Expanded Select Roles Drop-down Menu



6. From the *Select Organization(s)* drop-down lists, select organizations that will receive the file(s) you send (see Figure 47). These drop-down lists adhere to TIDE's user role hierarchy. For example, district-level users will be able to filter at their role level and below.

Figure 47. Inbox Window: Send Files Tab with Select Organization(s) Panel in Use



7. If you selected **By Role** in step 3, skip this step. If you selected **By Email** in step 3, enter the email address of the recipient to whom you wish to send a file or files.

8. To select a file or files to send, in the *Add File* field, select **Browse**. A file browser appears.

9. Select the file(s) you wish to send. You may send up to 10 files totaling no more than 20MB at once.

10. Select **Send**.

T

Test Opportunity Status Descriptions

You can view descriptions of each status in the table below when you view reports of test status codes.

Status	Definitions
Approved	The TA has approved the student for the session, but the student has not yet started or resumed the test.
Completed	The student has submitted the test for scoring. No additional action can be taken by the student.
Denied	The TA denied the student entry into the session. If the student attempts to enter the session again, this status will change to “Pending” until the TA approves or denies the student.
Expired	The student’s test has not been completed and cannot be resumed because the test has expired.
Invalidated	The test result has been invalidated.

Status	Definitions
Paused	The student's test is currently paused (as a result of one of the following): • The student paused his or her test by selecting the **Pause** button. • The student idled for too long (more than 20 minutes) and the test was automatically paused. • The test administrator-level user stopped the session the student was testing in. • The test administrator-level user paused the individual student's test. The student's browser or computer shut down or crashed.
Pending	The student is awaiting TA approval for a new test opportunity.
Reported	The student's score for the completed test in TDS has passed the quality assurance review and has been submitted to Reporting. Some items must be hand scored before they appear in Reporting.
Review	The student has answered all test items and is currently reviewing his or her answers before submitting the test. (A test with a "review" status is not considered complete.)
Scored	The test will display a scored status, followed by the student's score.
Started	The student has started the test and is actively testing.
Submitted	The test has been submitted for quality assurance review and scoring before it is sent to Reporting. Note: All tests go through an internal scoring process during quality assurance review.

U

User Role Permissions

Each user in TIDE has a role, such as a district-level user or a test administrator-level user. Each role has an associated list of permissions to access certain features within TIDE. Refer to the User Roles and Access document on the Montana testing portal for information on which users can access specific features and tasks within each CAI system. User guides are available for each system containing complete information about system features and functions.

User Support

For additional information and assistance in using TIDE, contact the CAI Helpdesk.

The Helpdesk is open 6:00 a.m. – 6:00 p.m. MT (except holidays or as otherwise indicated on the State Assessment Portal).

Montana Assessment Helpdesk

Toll-Free Phone Support: 1-877-365-7915

Email Support: mthelpdesk@cambiumassessment.com

Please provide the Helpdesk with a detailed description of your problem, as well as the following:

- If the issue pertains to a student, provide the SSID and associated district or school for that student. Do not provide the student's name.
- If the issue pertains to a TIDE user, provide the user's full name and email address.
- Any error messages that appeared.
- Operating system and browser information, including version numbers (e.g., Windows 7 and Firefox 13 or Mac OS 10.7 and Safari 5).

V

Valid Values for Tool Names in the Test Settings Upload File

Tool Name	Description	Valid Value	Applies to
Braille Transcript	Availability of braille transcripts.	Off	ELA
		On	ELA
Braille Type	Type of Braille in which test items are printed.	UEB Contracted with Nemeth Math	ELA, ELA-PT, Mathematics, Science
		UEB Uncontracted with Nemeth Math	ELA, ELA-PT, Mathematics
		UEB Contracted with UEB Math	Mathematics
		UEB Uncontracted with UEB Math	Mathematics
		Not Applicable	ELA, ELA-PT, Mathematics, Science
Color Choices	Selected color of test text and background.	Blue, Light blue, Black on cream, Gray, Light gray, Medium Gray on Light Gray Green, Light green, Magenta, Light magenta, White on navy, White on red, Red on white, Yellow, Light yellow, Yellow on Blue; Yellow on black, and Reverse Contrast.	ELA, ELA-PT, Mathematics, Science, Alt Science

Tool Name	Description	Valid Value	Applies to
Closed Captioning (Listening Items)	Availability of closed captioning.	Off	ELA
		On	ELA
Glossaries (English and Translated)	List of available glossaries.	English	ELA, ELA-PT, Mathematics
		Arabic Glossary, Burmese Glossary, Cantonese Glossary, Filipino Glossary, Hmong Glossary, Illustration Glossary, Korean Glossary, Mandarin Glossary, Punjabi Glossary, Russian Glossary, Somali Glossary, Spanish Glossary, Tagal Glossary, Ukrainian Glossary, Vietnamese Glossary All translated language glossaries can be combined with English and Illustration.	Mathematics
		No Word List Available	ELA, ELA-PT, Mathematics
Language	Sets language display for test content.	English	ELA, ELA-PT, Mathematics, Science, Alt Science
		Braille	ELA, ELA-PT, Mathematics, Science
		Spanish	Mathematics, Science
Masking	Availability of masking tool.	Off	ELA, ELA-PT, Mathematics, Science, Alt Science
		On	ELA, ELA-PT, Mathematics, Science, Alt Science
Mouse Pointer [Size and Color]	List of available mouse pointers.	Large Black, Extra Large Black, Large Green, Extra Large Green, Large Red, Extra Large Red, Large White, Extra Large White, Large Yellow, Extra Large Yellow	ELA, ELA-PT, Mathematics, Science, Alt Science

Tool Name	Description	Valid Value	Applies to
		System Default	ELA, ELA-PT, Mathematics, Science, Alt Science
Permissive Mode	Availability of permissive mode.	Off	ELA, ELA-PT, Mathematics, Science, Alt Science
		On	ELA, ELA-PT, Mathematics, Science, Alt Science
Print on Request	Sets which test content can be requested for printing.	Items	ELA, ELA-PT, Mathematics, Science, Alt Science
		Passages	ELA
		Passages and Items	ELA
		Stimuli	ELA-PT, Mathematics, Science, Alt Science
		Stimuli and Items	ELA-PT, Mathematics, Science, Alt Science
		None	ELA, ELA-PT, Mathematics, Science, Alt Science
Streamlined Mode	Availability of streamlined mode.	Off	ELA, ELA-PT, Mathematics, Science, Alt Science
		On	ELA, ELA-PT, Mathematics, Science, Alt Science
Sign Language (Embedded)	Availability of American Sign Language video.	Off	ELA, Mathematics
		On	ELA, Mathematics
Speech-to-Text (Embedded)	Availability of Speech-to-Text	Off	ELA, ELA-PT, Mathematics
		On	ELA, ELA-PT, Mathematics
Text-to-Speech (Designated Supports and Accommodations)	Sets which test content should read aloud.	Items	ELA, ELA-PT, Mathematics, Science
		Passages	ELA
		Passages and Items	ELA
		Stimuli	ELA-PT, Mathematics, Science

Tool Name	Description	Valid Value	Applies to
		Stimuli and Items	ELA-PT, Mathematics, Science
		None	ELA, ELA-PT, Mathematics, Science
Zoom	Sets default zoom level for test content.	1X	ELA, ELA-PT, Mathematics, Science, Alt Science
		1.5X, 1.75X, 2.X, 3X, 5X, 10X, 15X, 20X	ELA, ELA-PT, Mathematics, Science, Alt Science

Appendix 5-D
Reporting System User Guide

Montana Office of Public Instruction

Reporting System User Guide

For Summative and Interim Assessments

2022–2023

Published September 7, 2022

Prepared by Cambium Assessment, Inc.

Table of Contents

Introduction to the User Guide

This user guide gives instructions on using the Reporting System for the following:

- Accessing summative assessment data.
- Accessing interim assessment data.
- Scoring interim assessments.

It includes the following sections:

- How to Navigate Reports for Summatives and Interims
- How to Set Up Reports for Summatives and Interims to Suit Your Needs
- How to Export and Print Data on Summatives

The guide also includes additional information on accessing your interim data and scoring interims:

- How to Access Item-Level Data on Interims
- How to Score Items on Interims
- How to Set Up Interim Reports to Suit Your Needs
- How to Export and Print Data on Interims

How to Navigate Reports for Summatives and Interims

This section explains how to navigate your reports for both summative and interim assessments.

How to Understand Which Students Appear in Your Reports

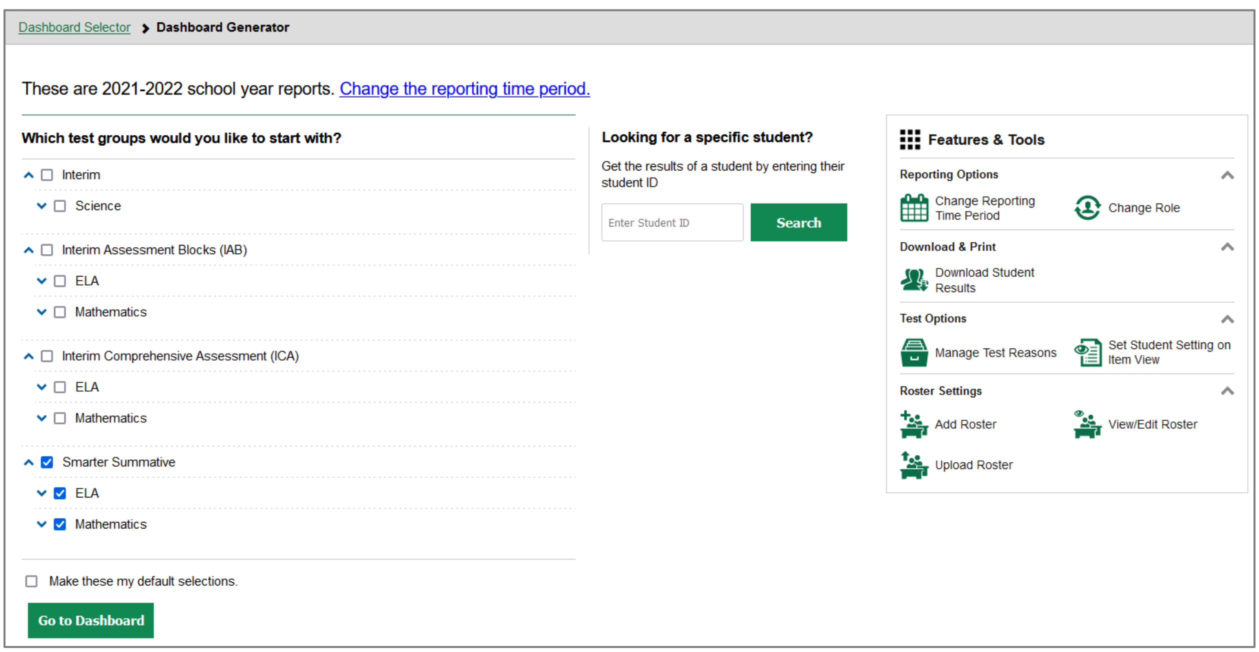
- Teachers can view data for all students in their classes (rosters) who have completed assessments. They can also view data for students to whom they have administered assessments in the current school year.
- School-level users can view data for all students in their schools who have completed assessments.
- District-level users can view data for all students in their districts who have completed assessments.

How to Start Viewing Reports

How to Use the Dashboard Generator Page

When you log in to the Reporting System, the Dashboard Generator page appears (Figure 1). The controls on the left let you select from the groups of tests that have been processed in your state for the school year. Typically, you may choose test types, subjects within the test types, and grades within the subjects.

Figure 1. Dashboard Generator Page



To the right of the dashboard generation controls are some other useful features. The *Looking for a specific student?* section allows you to use student ID to view all that student's test results. The **Features & Tools** menu appears on the right side of this page and in reports and offers multiple features that are described elsewhere in this guide.

If the Dashboard Generator shows a message saying there are no data, that means no test opportunities are available for this school year.

1. *Optional*: To change the test groups listed in the Dashboard Generator, change the reporting time period.
2. Select the tests you want to view, expanding the test groups as needed.
3. *Optional*: To save your selections, mark the checkbox **Make these my default selections**. These selections will be set in the Dashboard Generator whenever you log in. You can change the defaults at any time.
4. Click **Go to Dashboard**. The dashboard appears, displaying any data available for your selections.

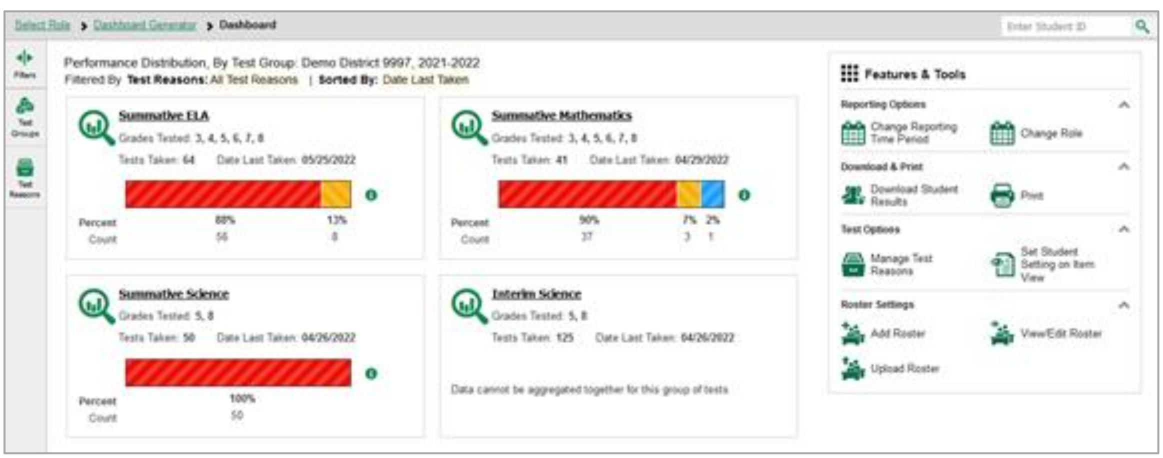
How to Use the Dashboard to View Aggregate Test Results

All users can view the standard dashboard (Figure 2). It displays aggregation cards representing test groups.

Each aggregation card displays the test group name, a list of grades, the number of students who took tests in the group, the date of the test last taken, and a performance distribution bar displaying both percentages and student counts below it. You may sometimes see the message "Data cannot be aggregated together for this group of tests" instead of the performance distribution bar for tests that do not report performance distribution, or that use different sets of performance levels. Test group cards are sorted by date last taken.

Clicking the ℹ button beside the performance distribution bar displays a legend with more information about performance levels.

Figure 2. Dashboard



If a message appears saying "There are no assessments to display," you may not have any students who have taken tests in your selected test groups in the selected time period. You can change the reporting time period. If you are a teacher, you may also be able to view more students' data by managing your classes (rosters).

To change the test groups and test reasons that appear, use the **Filters** panel on the left. For more information on filtering, see How to Set Up Reports for Summatives and Interims to Suit Your Needs and How to Set Up Interim Reports to Suit Your Needs. You can also backtrack to the Dashboard Generator using the link in the path at the upper-left corner and change your test group selections there.

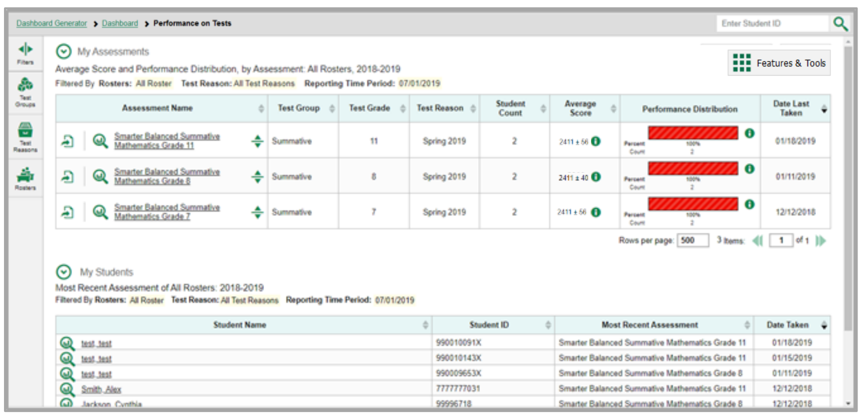
How to View More Detailed Data on a Particular Test Group

To view more detailed data for a particular test group, click the name of the group (or beside it). The Performance on Tests report appears. It is filtered to display only the test group you selected.

In the Performance on Tests report, teachers see two tables, as in Figure 3:

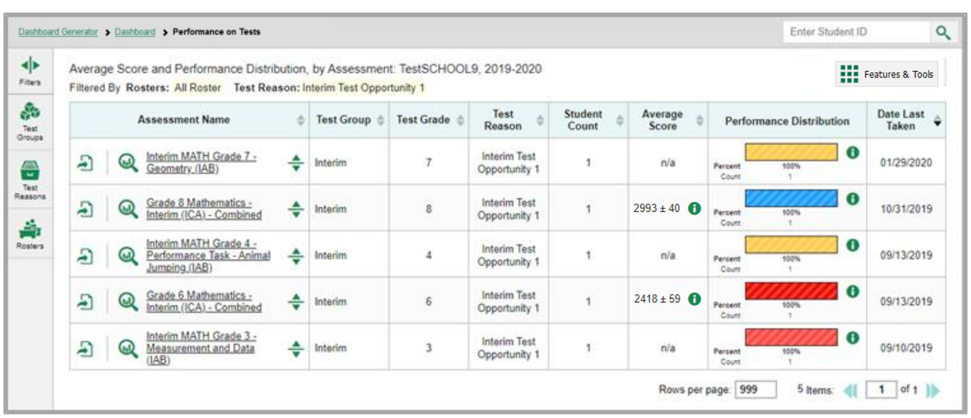
- The My Assessments table, listing all the assessments in the filtered test group or groups.
- The My Students table, listing all your students who took the assessments.

Figure 3. Teacher View: Performance on Tests Report



District- and school-level users see just one table, as in Figure 4. Like the first table on the teacher Performance on Tests report, this table lists all the assessments in the test group.

Figure 4. School-Level User View: Performance on Tests Report



For each test, the assessments table shows the test group, grade, test reason (the name of the test window of a summative assessment, or a category assigned to an interim assessment), number of students who took the test, average score, performance distribution, and date the test was last taken.

You can use the filters to view a different set of assessments. For more information on filtering, see How to Set Up Reports for Summatives and Interims to Suit Your Needs and How to Set Up Interim Reports to Suit Your Needs.

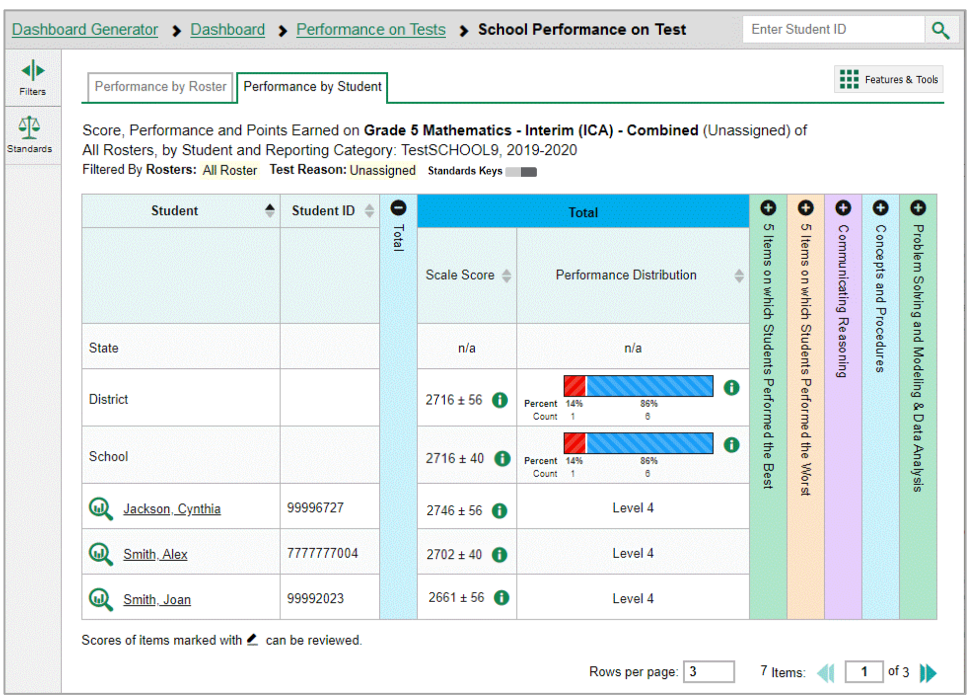
If a message appears saying “There are no assessments to display” or “There are no students to display”, you may have filtered out all data.

For Teachers and School-Level Users: How to View Student-Level Data for All Your Students

The **Performance by Student** tab displays test results for each of your students across classes (rosters). In order to see the results for all your students, follow the instructions below.

1. Starting from the dashboard that appears when you log in, click a test group name (or beside it).
2. Click a test name (or beside it) in the assessments table at the top of the page.
3. In the report that appears, select the **Performance by Student** tab, as in Figure 5. You will see results listing all your students. The first few rows also show aggregate performance data for your state, district, school, and/or total students.

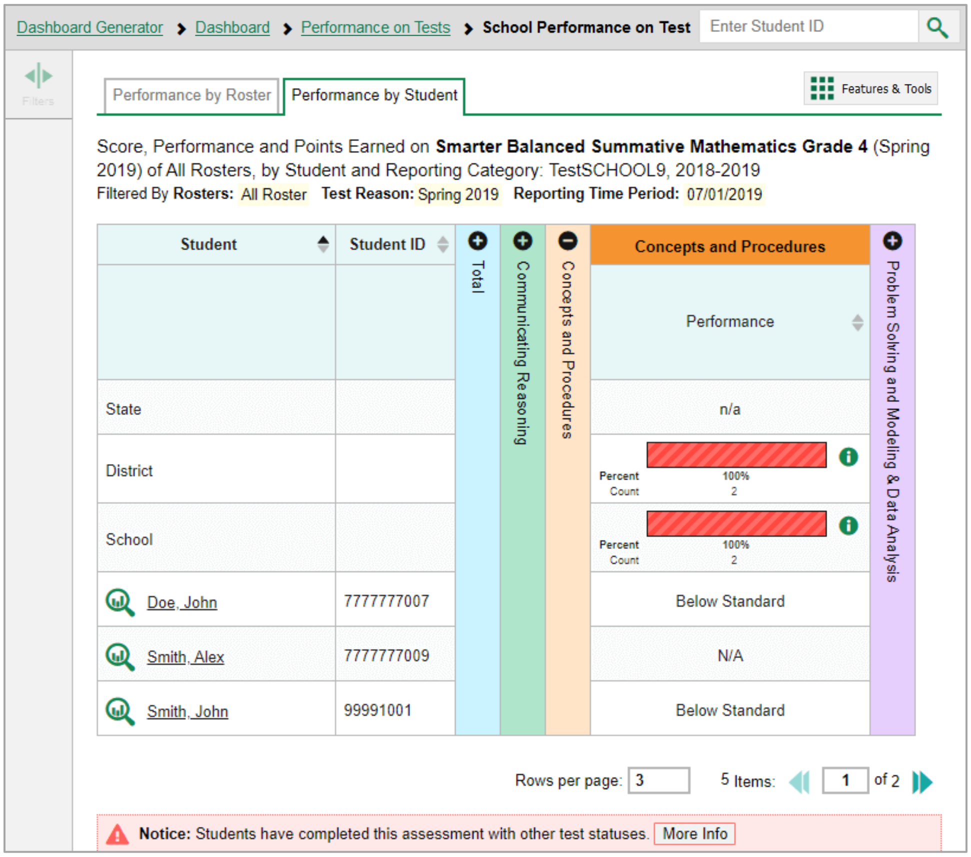
Figure 5. School Performance on Test Report: Performance by Student Tab



To see which students performed best, click the score or Performance columns to sort them.

You can view your students' performance in each area of the test by clicking the reporting category section bars to expand them, as shown in Figure 6.

Figure 6. My Students' Performance on Test Report: Performance by Student Tab with Expanded Reporting Category Section



How to View Test Results for Classes (Rosters) on a Particular Test

You can view a list of classes (rosters) that took a particular test, and you can also view the test results for a particular class.

How to Access Test Results for All Your Classes (Rosters)

The **Performance by Roster** tab (Figure 7) displays test results for each class (roster). To view this tab, follow the instructions for your user role below.

Teachers and school-level users:

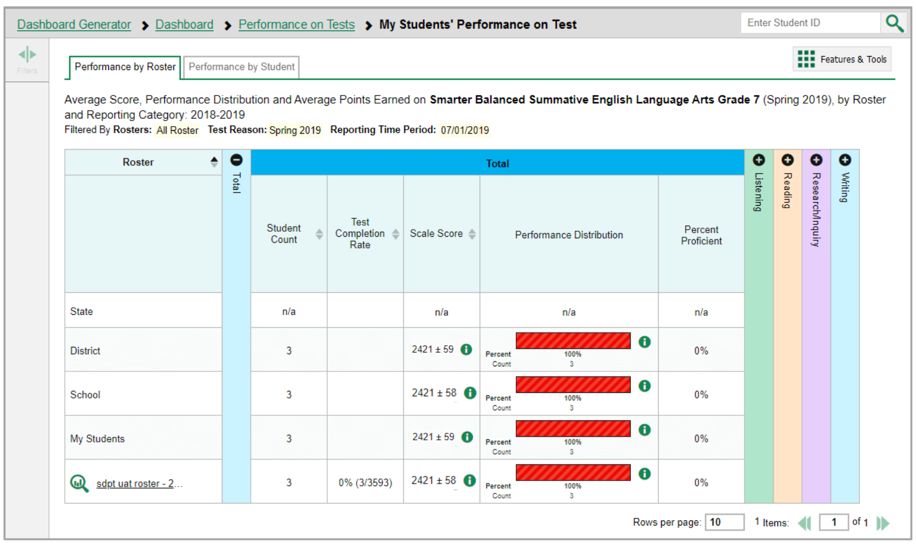
1. Generate a dashboard and click a test group name (or beside it).
2. Click a test name (or beside it) in the table of assessments. Either the My Students' Performance on Test or the School Performance on Test report appears, depending on your role. It is open to the **Performance by Roster** tab.

District-level users can view all classes (rosters) in a school. To do so, follow these instructions:

1. Generate a dashboard and click a test group name (or beside it).
2. Click a test name (or beside it) in the table of assessments. A page of district test results appears, listing schools within the district.
3. Click a school name (or beside it). The School Performance on Test report appears, open to the **Performance by Roster** tab.

The report shown in Figure 7 displays a list of your classes (rosters) and each class's performance. The first few rows also show aggregate performance data for your state, district, school, and total students.

Figure 7. My Students' Performance on Test Report: Performance by Roster Tab



How to See Which Classes (Rosters) Performed Well on This Assessment

To see which classes performed best on the test, do either of these things:

- Click the score column header to sort by score and look for rosters with high average scores.
- Look at the bars in the Performance Distribution column to see where the percentage of students at or above proficient is high.

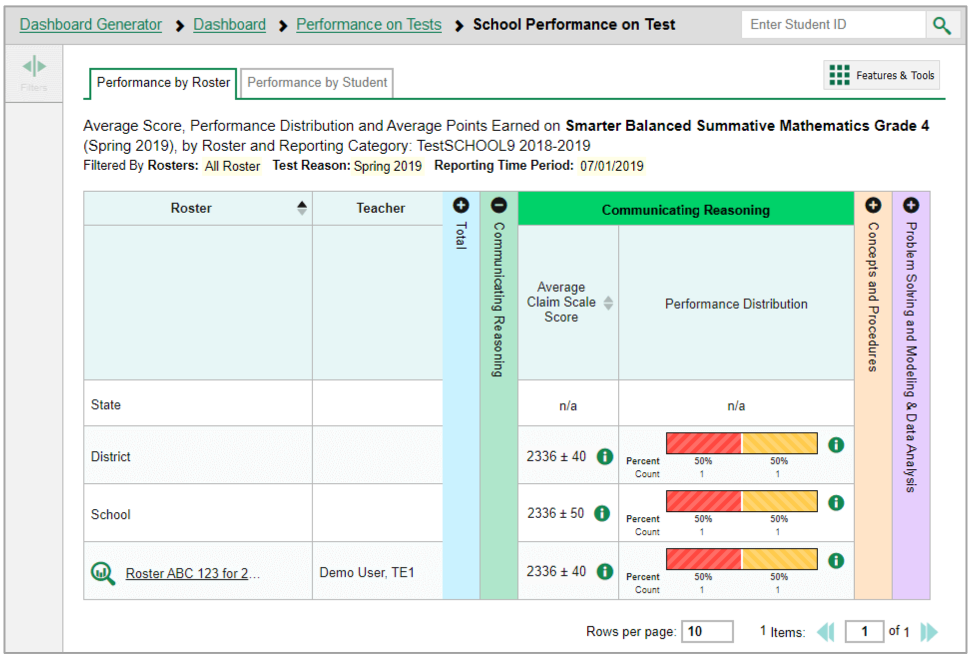
How to See Which Classes (Rosters) Had the Highest Test Completion Rates

To see which classes had the highest test completion rate, click the Test Completion Rate column header to sort the column.

How to See How Well Classes (Rosters) Performed in Each Area on the Test

For tests with reporting category sections, you can compare the performance of your students in each area of the test. Click each vertical section bar to expand or collapse it. In this example (Figure 8), you can view average score and a performance distribution bar for each class (roster) under the reporting category Communicating Reasoning.

Figure 8. My Students' Performance on Test Report: Performance by Roster Tab with Expanded Reporting Category Section



Summative and interim tests cover multiple reporting categories, while a benchmark test covers only one.

How to View and Interpret Standard Measures Within Reporting Category Sections

Aggregate test results for Smarter Balanced and Montana Science Assessment (MSA) tests may include detailed performance measures within standards within reporting categories.

An educational standard, sometimes called an assessment target, describes the skill the item measures. Standards are nested within clusters (groups of standards).

Note: The Alternate Montana Science Assessment (AMSA) tests are reported at the overall scale score only. No sub-scores or reporting categories are included.

The Standards sub-section (shown in Figure 9) contains the following:

- **Clusters** within the reporting category.
    - **Standards** within each cluster.

  - **Measures** within each standard.

To learn more about each standard, click the more information button to the right of the standard name.

Figure 9. School Performance on Test Report: Performance by Roster Tab with Expanded Reporting Category Section

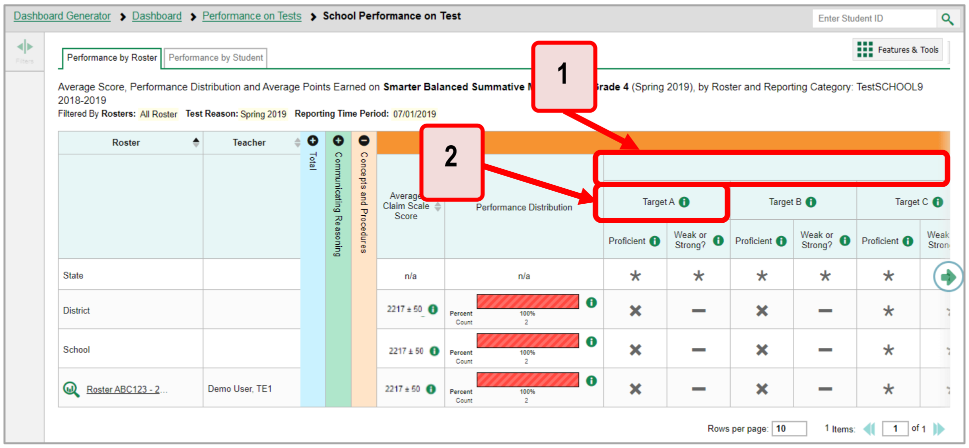


Table 1. School Performance on Test Report: Performance by Roster Tab: Reporting Category Section Elements

#	SBAC Elements	MSA Elements
1	Claim	Discipline
2	Targets	Disciplinary Core Idea (DCI)

Each standard column consists of sub-columns displaying performance measures:

- **Weak or Strong?:** You may want to identify strengths and weaknesses in different standards within the reporting category so you can focus instruction on very specific areas where improvement is needed. This column displays icons indicating how students' average performance on the standard compares with their average performance on the overall test. Note that these icons indicate only relative performance and not proficiency.
- **Proficient:** This column displays icons indicating whether students have on average attained proficiency in the standard.

To learn more about these measures and the symbols they use, click the more information button to the right of each measure.

How to View and Interpret Writing Dimension Measures (ELA ICA Only)

Aggregate test results for some assessments may also include a **Writing Dimensions** section to the right of the expandable sections in the report table. You can expand it by clicking the vertical bar, just as with the reporting category sections. This section helps you understand how students performed on different aspects of writing.

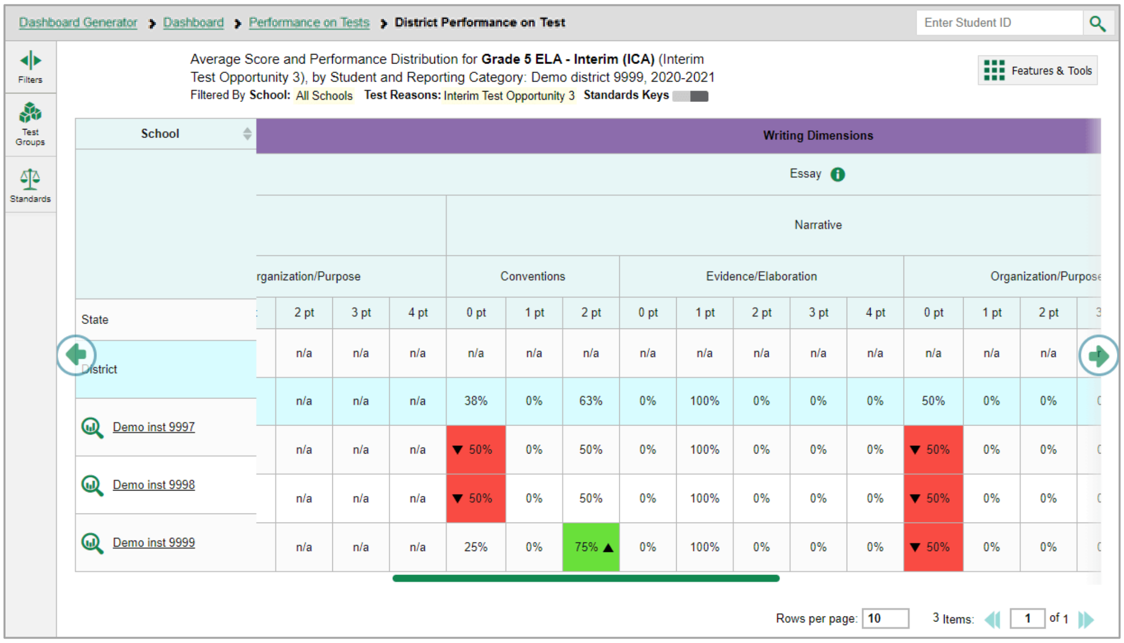
The **Writing Dimensions** section (see Figure 10) contains the following under the Essay header:

- **Essay type**. For example, Informative/Explanatory, Opinion, and Narrative.
  - **Writing dimension** within the essay type, as listed in item rubrics. For example, Purpose, Focus, and Organization; Evidence and Elaboration; and Conventions of Standard English.
    - **Point value**. A sub-column for each possible item point value for the writing dimension. Each point value sub-column displays the percentage of students who earned that number of points.

    For each dimension, the lowest ▼ and highest ▲ point values are sometimes highlighted and marked with arrow icons in the rows with the highest percentages. This allows you to quickly identify groups of students who are performing well and those who may need additional support.

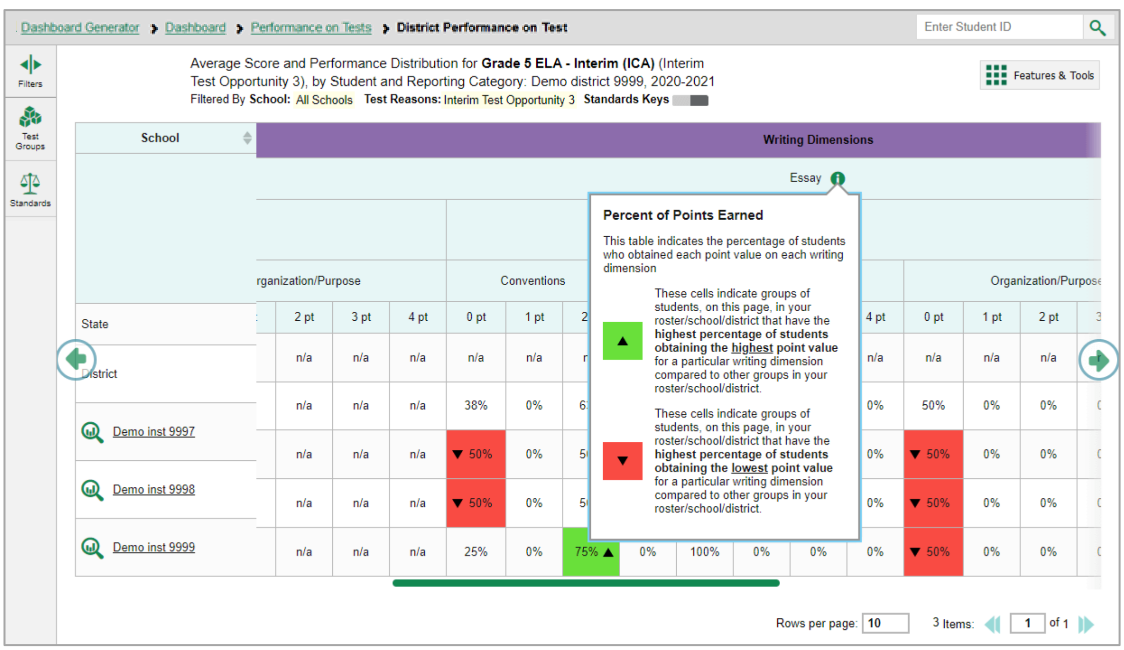
    Note that highlighting and arrow icons do not appear where they are not useful. A report containing only one row of data does not have them, and neither does a row in which the percentages are "n/a" or are all the same.

Figure 10. District Performance on Test Report with Expanded Writing Dimensions Section



To learn about the highlighting and arrow icons, click the more information button in the Essay header. A legend expands, as in Figure 11.

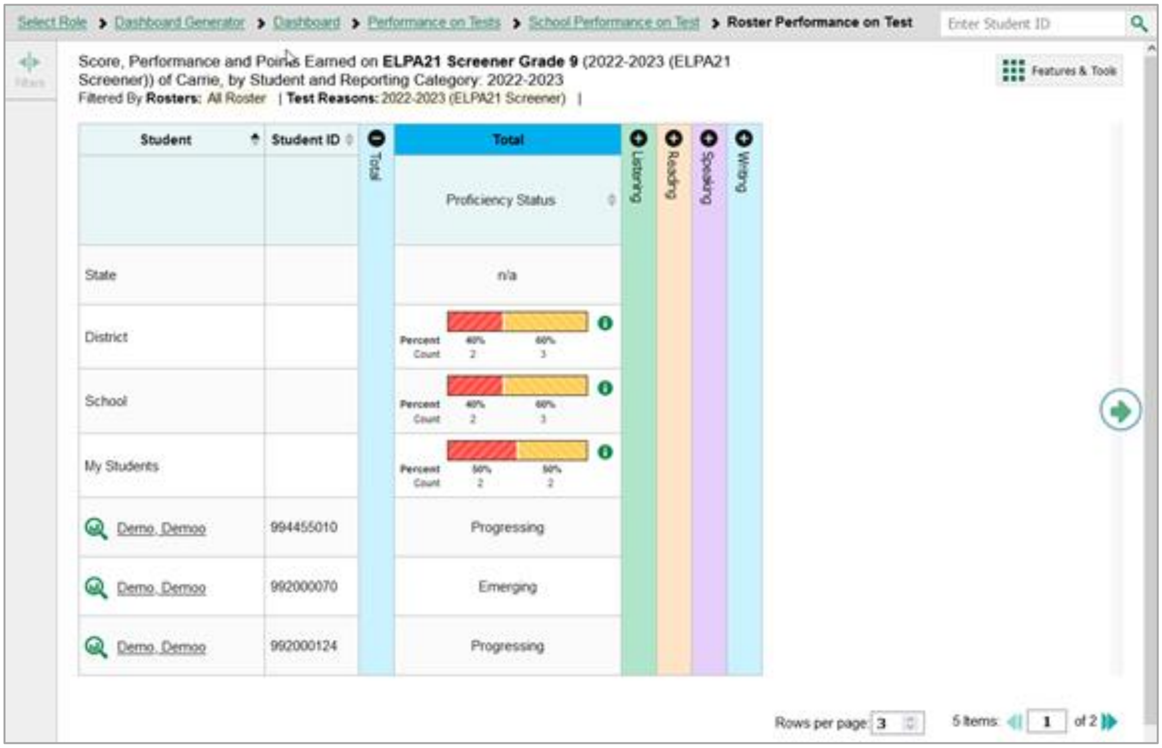
Figure 11. District Performance on Test Report with Expanded Writing Dimensions Section and Expanded Legend



How to Access Test Results for an Individual Class (Roster)

Prior sections explained how to access test results for all your classes (rosters). To view results for one specific class, click the name of a class that appears in the first column of the report (or beside it). The class results listed by student appear (see Figure 12).

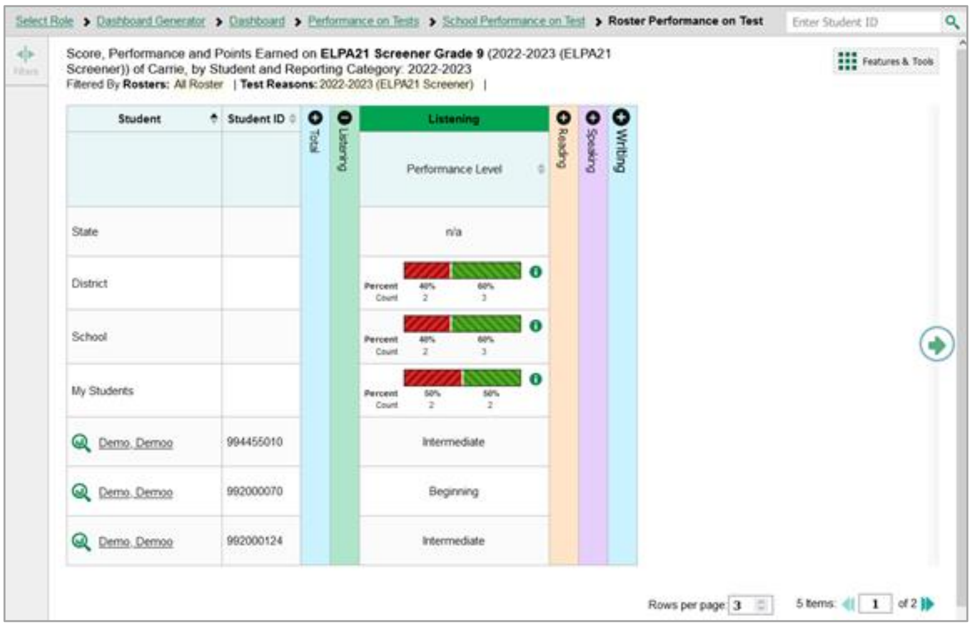
Figure 12. School-Level User View: Roster Performance on Test Report



How to See How Well Students in Your Class (Roster) Performed in Each Area of the Test

You can compare the performance of your students in each area of the test using the reporting category sections, as shown in Figure 13. Click the vertical section bar to expand each section.

Figure 13. School-Level User View: Roster Performance on Test Report with Expanded Reporting Category Section



For School- and District-Level Users: How to View Test Results for a School on a Particular Test

You can view test results for all the students in a school on a particular test. This gives you a high-level look at how the school is performing.

How to Access Test Results for a School

School-level users:

1. Generate a dashboard and click a test group name (or beside it).
2. Click a test name (or beside it) in the table of assessments. The School Performance on Test report appears.

District-level users:

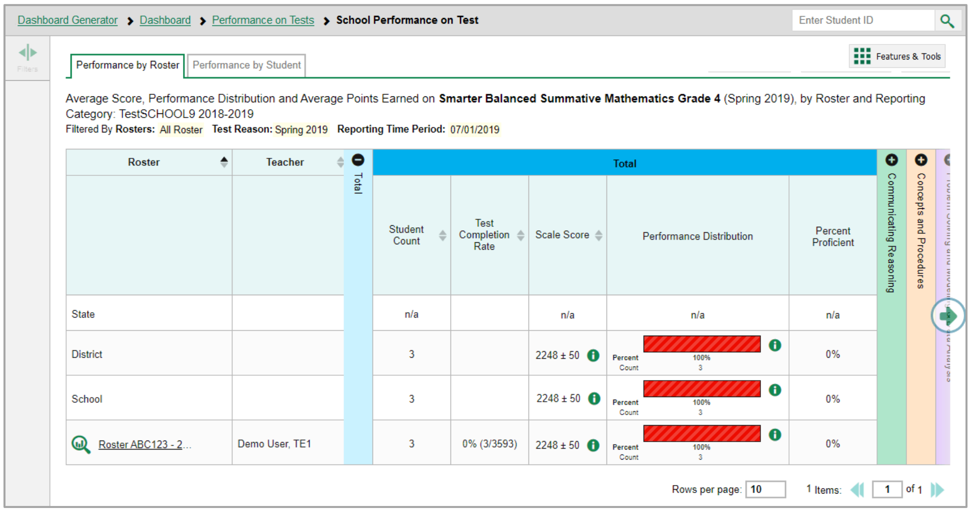
1. Generate a dashboard and click a test group name (or beside it).
2. Click a test name (or beside it) in the table of assessments. A table listing test results by school appears.

3. Click the name of the school (or beside it) for which you would like to see results. The test results for the school appear.

Note that district-level users can also access the test results from a school directly from the Performance on Tests report, by first filtering by school.

The **Performance by Roster** tab is open by default, as in Figure 14.

Figure 14. School Performance on Test Report: Performance by Roster Tab



How to See Which Classes (Rosters) Performed Well on This Assessment

In the **Performance by Roster** tab, look at the Performance Distribution column and click the header of the score column to sort by score. Rosters with a high average scale score, and with a high percentage of students performing at or above proficient in the performance distribution bar, performed well on the assessment. If certain classes (rosters) performed consistently well, you could use them as a model for the classes with lower performance.

For District-Level Users: How to View Test Results for a District on a Particular Test

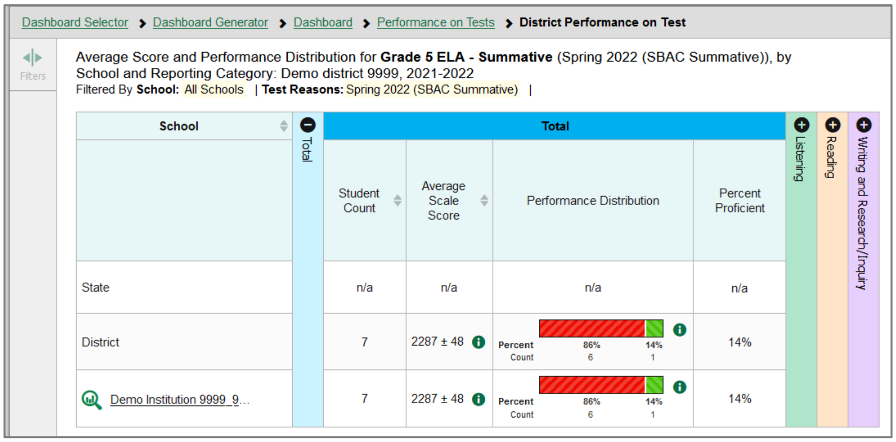
You can view test results for a district on a particular test. This gives you a high-level view of how the district is performing.

How to Access Test Results for a District

1. Generate a dashboard and click a test group name (or beside it).

2. Click a test name (or beside it) in the table of assessments. The District Performance on Test report appears, listing schools in the district (see Figure 15).

Figure 15. District Performance on Test Report



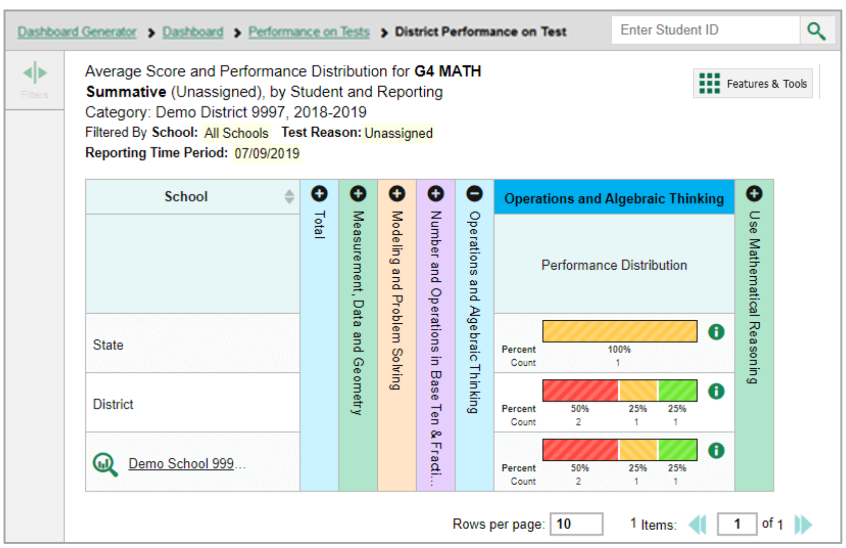
How to See Which Schools in the District Performed Well on This Assessment

Look at the score column and/or Performance Distribution column, and click the score column to sort by it.

How to See How Well Schools in the District Performed in Each Area of the Test

Click the vertical section bars to expand the reporting category sections (as in Figure 16).

Figure 16. District Performance on Test Report with Expanded Reporting Category Section



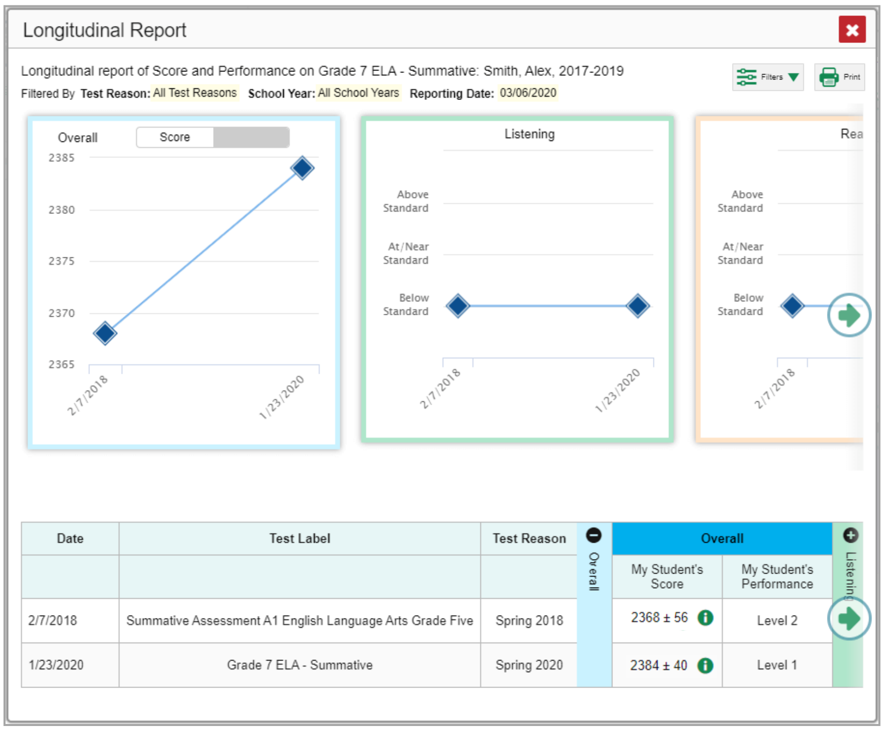
How to Track Student Performance Over Time

You can view your students' performance over time across multiple related assessments or across multiple test opportunities of a single assessment. This lets you see how students' performance has improved or declined.

Each Longitudinal Report displays performance data for one of the following:

- A group of students who have completed every assessment available in the report. If you are a district- or school-level user, note that a certain percentage of students must have taken *all* the related assessments in order for you to generate a Longitudinal Report. Teachers have the option of adjusting the pool of students, tests, and test reasons.
- An individual student (see Figure 17).

Figure 17. Longitudinal Report Window: Summative Report for a Single Student with Multiple Reporting Categories



How to Access a Longitudinal Report Comparing Related Assessments

If the student(s) in your test results have completed multiple related assessments, the **Build Longitudinal Report** button allows you to access a Longitudinal Report in the reports for any of those assessments. If they haven't done so, then no Longitudinal Report is available.

Click the **Build Longitudinal Report** button in the **Features & Tools** menu.

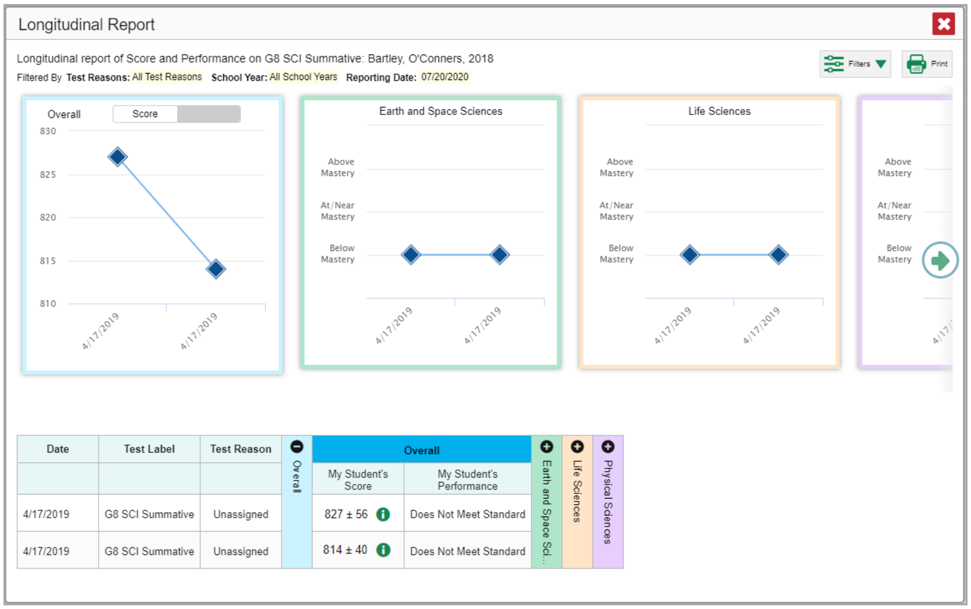
The ***Longitudinal Report*** window appears. Depending on your role, the test types, and the number of students in the report, it may display a report options page rather than the Longitudinal Report itself. The contents of this page depend on your role and the number of students.

How to View Students' Overall Performance on These Assessments Over Time

Look at the graph in the upper-left corner of the Longitudinal Report (see Figure 18). It shows the scores or performance levels of the student(s) each time they took the test(s).

Score data are plotted along a line.

Figure 18. Longitudinal Report Window: Summative Report for a Single Student with Multiple Reporting Categories



Performance level data are shown either the same way or, for multiple students, in performance distribution bars.

Mouse over the data points in a line graph or the sections in a bar to get more information.

Alternatively, in the table at the bottom of the report, look at the **Overall** section.

How to Switch Between Score Data and Performance Level Data

When a graph offers both score and performance level data, a toggle bar appears at the top of it. Click the toggle to switch. You may want to do this if you find performance level data easier to read, or if you prefer the precision of score data. Sometimes a test includes only one type of data.

How to See Students' Performance in Different Areas Over Time

Look at the reporting category graphs to the right of the overall performance graph, or look at the expandable reporting category sections in the table at the bottom. Here, you can see at a glance how students are improving or declining in each area, and you can compare their trajectories in different areas.

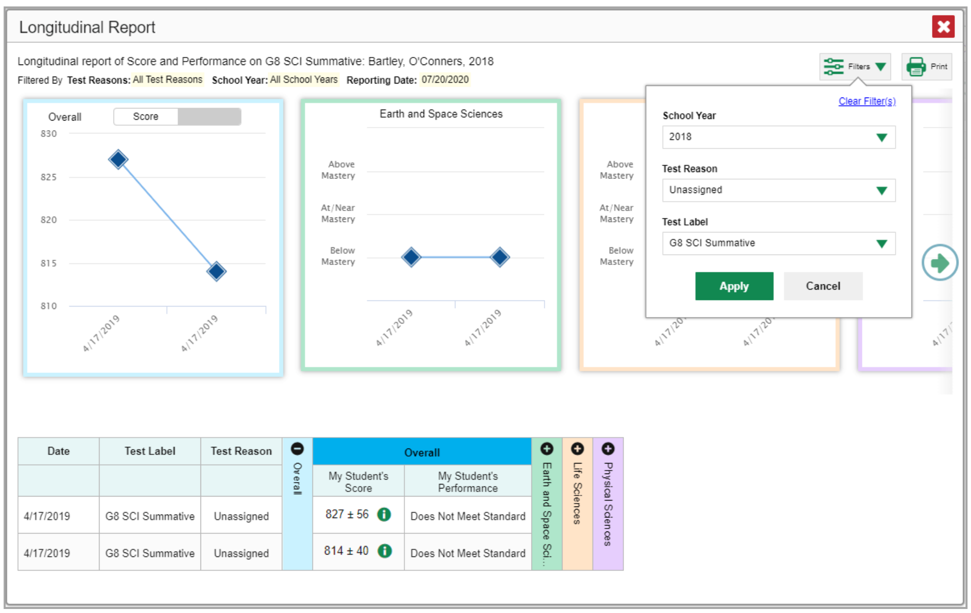
How to Filter Test Opportunities to Show Some and Not Others

You may want to filter a Longitudinal Report in order to focus on some test opportunities and not others.

Note that filtering tests may affect the set of students whose data are included in the report.

1. Open the **Filters** menu at the upper-right corner and select the filter options you prefer from the drop-down lists (see Figure 19).

Figure 19. Longitudinal Report Window: Summative Report for a Single Student with Multiple Reporting Categories and with Expanded Filter Menu



- You may want to filter by a particular school year or years. Note that years are not calendar years. “2021” refers to the 2021–2022 school year. By default, Longitudinal Reports show data for all years.

  Longitudinal Reports can show student performance from a time when the students were not yet associated with you. For example, if you are a seventh-grade teacher, you can use these reports to view your current students’ performance on last year’s sixth-grade tests.

- If the report includes interim assessments, you may wish to filter by a test reason (a category of test), which means excluding all other test reasons from the data. For example, you may want to narrow the report down to show only tests taken in the spring. For summative assessments, test reasons are the same as test windows and are not useful.

- Finally, you may find that certain individual tests are less relevant than others. In that case, you can use the **Test Label** options to deselect the names of the tests you don’t want to see.

2. Click **Apply**.

3. *Optional*: To revert all filters to their defaults, open the **Filters** menu again and click **Clear Filters**. Click **Apply**.

A row of filter details appears below the report header, showing the test reasons and school years included in the report.

How to View Test Results Broken Down by Demographic Sub-Groups

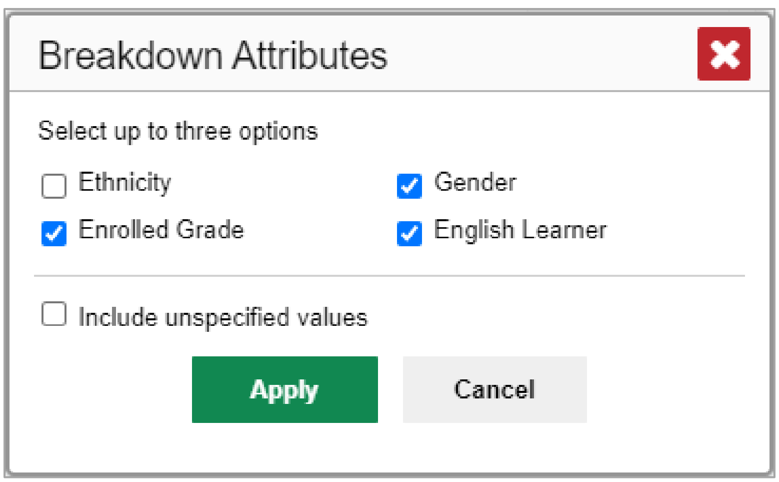
You can use the **Breakdown By** button in the **Features & Tools** menu to compare performance between different demographic sub-groups. This button is available for most aggregate test results.

How to View Test Results Broken Down by Demographic Sub-Groups

To view test results broken down by demographic sub-groups, do the following:

1. Click **Breakdown By** in the **Features & Tools** menu. The ***Breakdown Attributes*** window opens (see Figure 20).

Figure 20. My Students' Performance on Test Report: Performance by Student Tab: Breakdown Attributes Window



2. Select up to three student demographic categories.

   There is also an **Include unspecified values** checkbox, explained below:

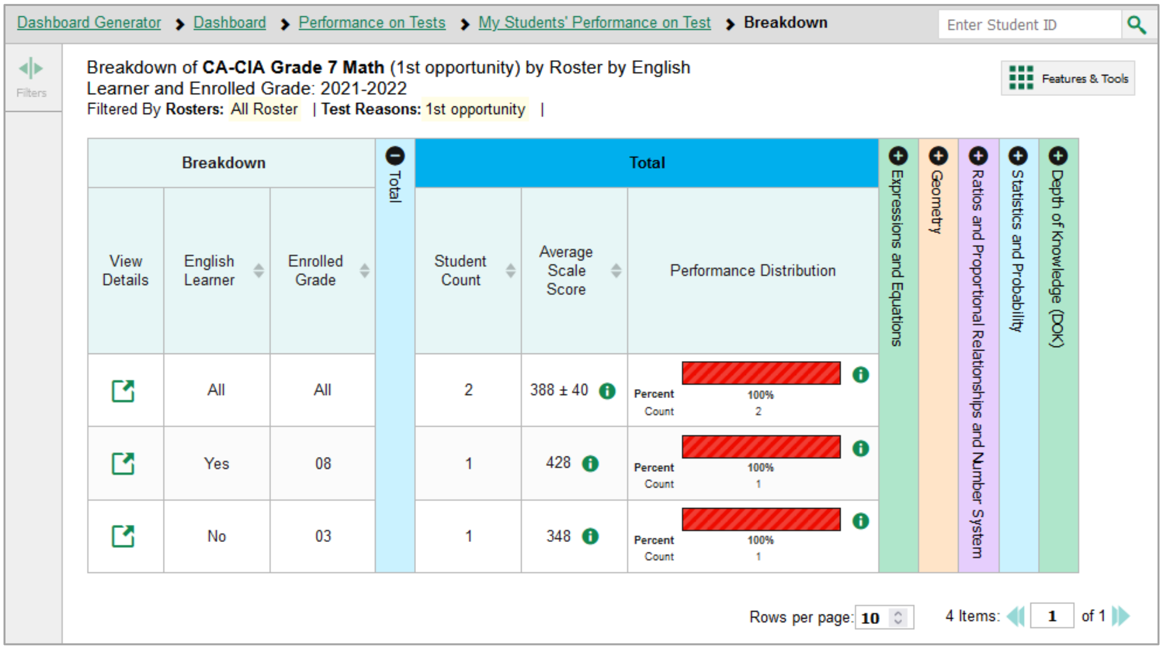
   - Some students who complete tests do not have specific demographic information in the Test Information Distribution Engine (TIDE). These students are considered to have unspecified values.

   - To include data for these students, mark the checkbox.

3. Click **Apply**.

Data for each sub-group selected are displayed in the report (see Figure 21).

Figure 21. Demographic Breakdown of a My Students' Performance on Test Report

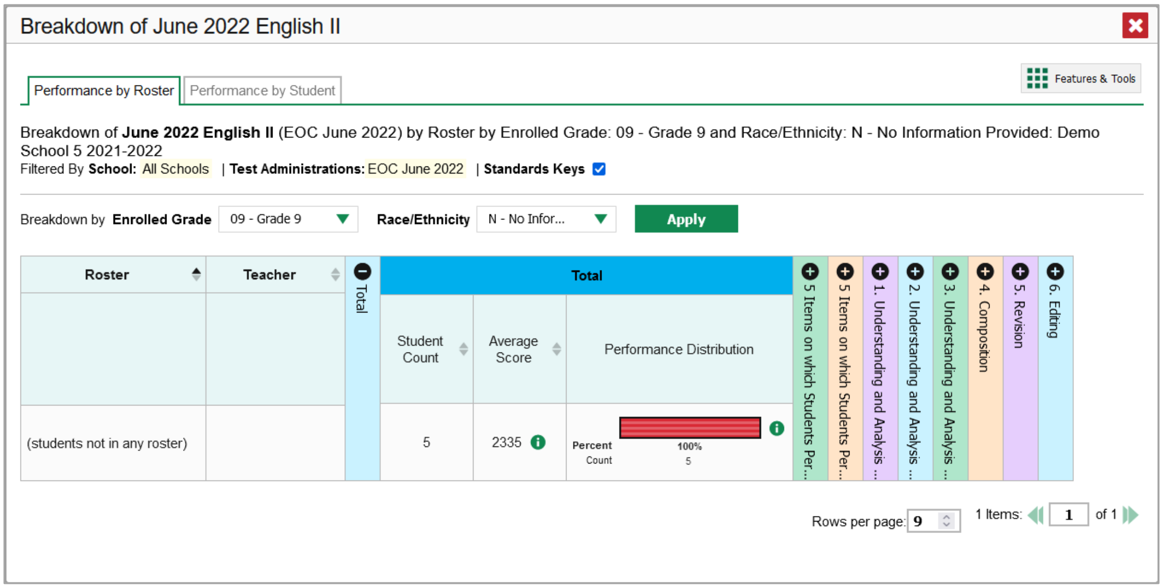


How to View Test Results for a Particular Demographic Sub-Group or Combination

When viewing test results broken down by sub-groups, go to the row for the demographic combination you want to view and click the view button in the View Details column on the left (see Figure 21).

A window opens, displaying detailed results for that combination. The report table is now laid out the same way as the original report, before you viewed it broken down by sub-groups (see Figure 22).

Figure 22. Demographic Combination Breakdown Window (from School Performance on Test Report)



At the top of the report table are filter menus for each demographic category you chose. To change the demographic combination displayed, use the filters to select the demographic sub-groups you want to see and click **Apply**. The new combination is displayed.

You can use this window to get an in-depth look at specific groups of students. For example, you may want to determine which classes (rosters) have the highest-performing girls in the first grade.

How to View Test Results for Individual Students

You can find out how well an individual student understands the material covered on a specific completed assessment. You can also view a report for all the assessments a student has taken.

How to Access Test Results for an Individual Student on a Particular Test

Teachers and school-level users:

1. Generate a dashboard and click a test group name (or beside it).
2. Click a test name (or beside it) in the table of assessments. A page of test results appears.
3. Select the **Performance by Student** tab.
4. Click the name of an individual student (or beside it) in the report. The Student Performance on Test report appears (see Figure 23).

District-level users:

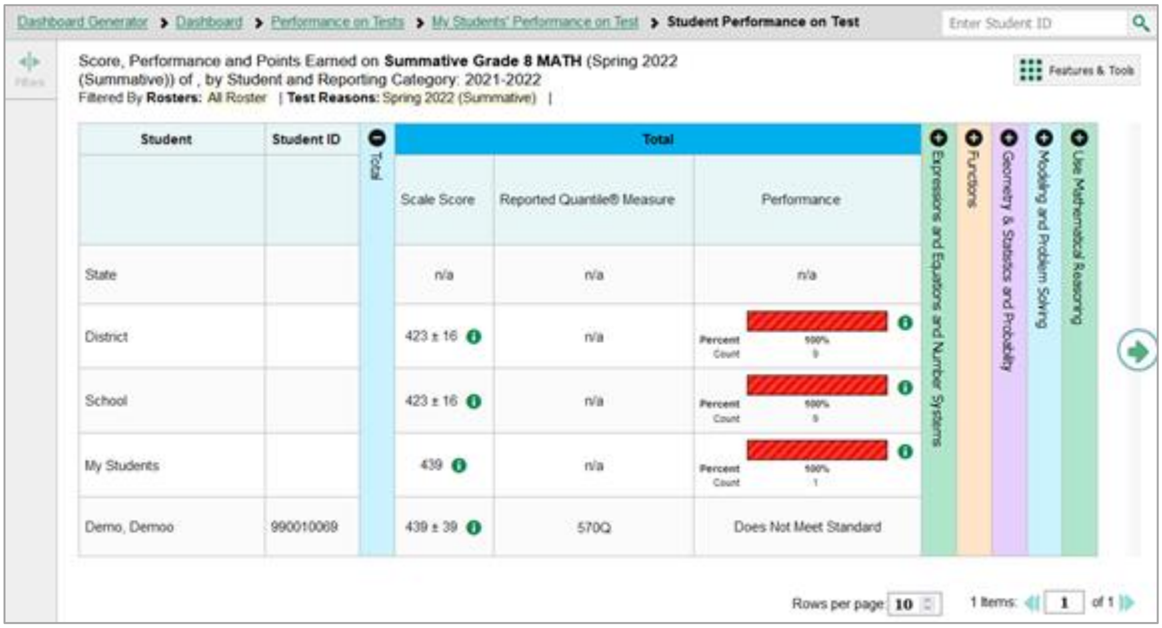
1. Generate a dashboard and click a test group name (or beside it).

2. Click a test name (or beside it) in the table of assessments. A page of test results by school appears.

3. Click a school name (or beside it). The School Performance on Test report appears.

4. Perform the same steps as teachers and school-level users, starting at step 3.

Figure 23. Teacher View: Student Performance on Test Report



You can view the student's performance in each area of the test using the reporting category sections, which you can click to expand.

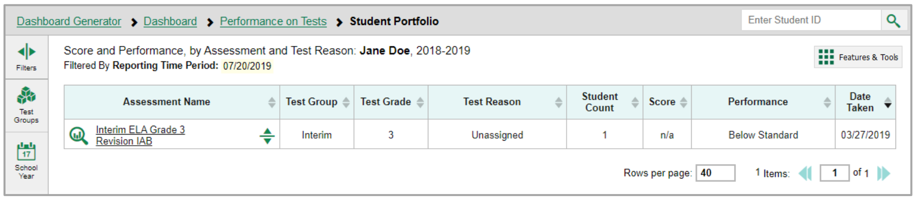
How to View a Report for All the Assessments a Student Has Completed Over Time

The Student Portfolio Report allows you to view all the assessments an individual student has completed over time. This is useful for viewing performance on tests that have multiple opportunities, and for interim tests that were administered multiple times throughout the year.

To access this report, enter the student's SSID in the search field in the upper-right corner and click . (When you are viewing the Dashboard Generator and data are available, the same field appears to the right of the dashboard generation controls.) The Student Portfolio Report appears (see Figure 24).

Teachers can also access this report from the Performance on Tests report by going to the My Students table below the main assessments table and clicking a student's name (or beside it).

Figure 24. Student Portfolio Report



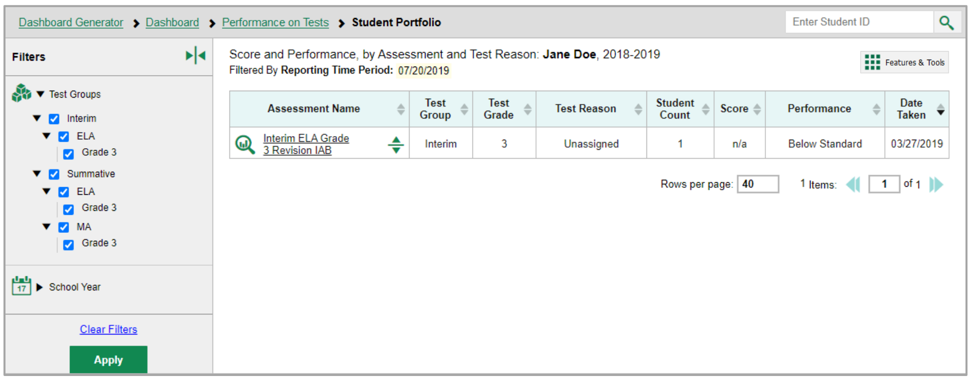
To view individual test results for this student, click a test name (or beside it).

How to Use the Student Portfolio Report to View Only the Tests You're Interested In

You can temporarily filter which tests you want to see in the Student Portfolio Report. You may want to do this, for example, if you are an ELA teacher and you don't want to see a student's math scores. By default, the data for those math assessments appear in the report, but you can exclude them.

1. In the **Filters** panel on the left side of the Student Portfolio Report, click either the expand button or the **Test Group** button . The **Filters** panel expands (see Figure 25).

Figure 25. Student Portfolio Report with Expanded Filters Panel



2. Mark as many selections as you like in the **Test Group** section of the **Filters** panel. Tests are organized by test type, subject, and grade.

3. Click **Apply**. The Student Portfolio Report updates to show only data for those tests.

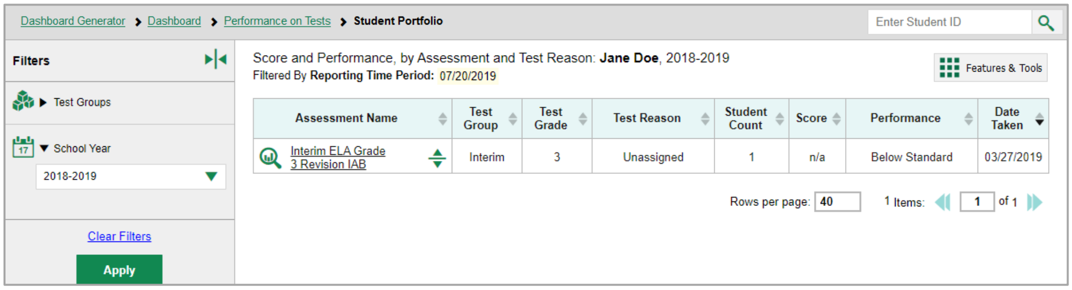
4. *Optional*: To revert all filters to their defaults, open the filters panel again and click **Clear Filters**. Click **Apply**. Filters also revert when you log out, switch user roles, or switch systems.

How to View a Student's Performance on Tests Taken in a Previous School Year or Years

If there are multiple years of data for a student, the Student Portfolio Report can look back at previous years. This gives you a high-level look at the student's progress. Student Portfolio Reports can show student performance from a time when the students were not yet associated with you. For example, if you are a seventh-grade teacher, you can use these reports to view a current student's performance on last year's sixth-grade tests.

1. On the left side of the page, click either the **Filters** panel expand button or the school year button . The **Filters** panel expands.
2. Under **School Year**, select a year or years (see Figure 26).

Figure 26. Student Portfolio Report with Expanded Filters Panel



3. Click **Apply**.

To switch back to the current year:

1. Open the **Filters** panel again.
2. Click **Clear Filters**.
3. Click **Apply**.

How to Use Individual Student Reports (ISR)

This section discusses Individual Student Reports (ISR), each of which provides easy-to-read performance data on a student's test. If a student took tests multiple times, an ISR will be available for each test opportunity (an instance of a test the student took). ISRs are useful for sharing performance information with students and their parents and guardians, and may be generated in the language of your choice.

What an Individual Student Report (ISR) Looks Like and How to Read It

An ISR is a PDF that shows results for a test opportunity. It may consist of a single page or multiple pages. ISR layouts vary according to the type of test. Details of sample ISRs are shown below in Figure 27, Figure 28, Figure 29, Figure 30, and Figure 31.

- At the top of each ISR are the student name and SSID, the name of the test, district, and school, and any other relevant information.
- Below that is a summary of the student's performance. An ISR for a scale-scored test displays the student's performance on a vertical scale that includes all the cut scores and performance levels.
- Each ISR includes a comparison table showing the average performance of the state, district, and/or school.
- Many ISRs include a table detailing the student's performance in each reporting category (as in Figure 28).
- Some ISRs include item-level data (as in Figure 29), scoring assertions, and/or scoring assertion outcomes (as in Figure 30).
- Some ISRs include longitudinal graphs (as in Figure 31).

Figure 27. Detail of Individual Student Report (ISR): Math Summative

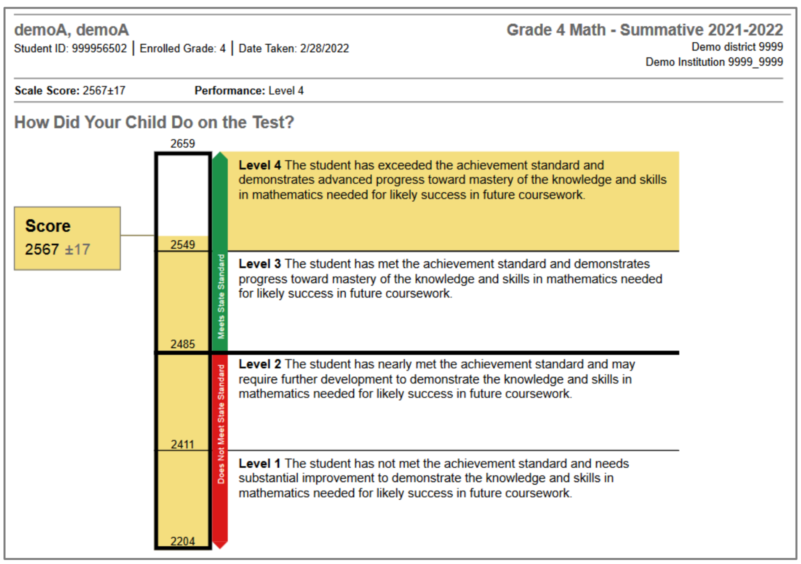


Figure 28. Detail of Individual Student Report (ISR): Math Interim (ICA) with Reporting Categories

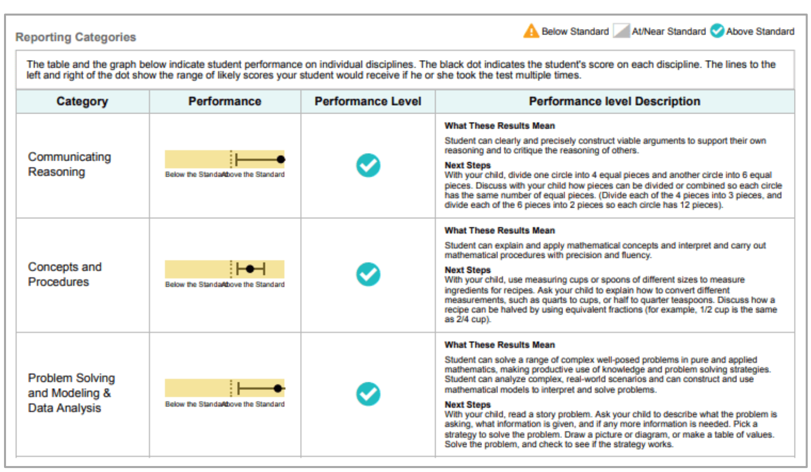


Figure 29. Detail of Individual Student Report (ISR): Math Interim (ICA) with Item- and Standard-Level Data

Communicating Reasoning		
Question #	**Standard**	**Points Earned/Points Possible**
3	Base arguments on concrete referents such as objects, drawings, diagrams, and actions.	0/2
5	Distinguish correct logic or reasoning from that which is flawed, and—if there is a flaw in the argument—explain what it is.	1/1
12	Test propositions or conjectures with specific examples.	0/1
17	State logical assumptions being used.	1/1
18	Use the technique of breaking an argument into cases.	1/1
28	Distinguish correct logic or reasoning from that which is flawed, and—if there is a flaw in the argument—explain what it is.	0/1
36	Base arguments on concrete referents such as objects, drawings, diagrams, and actions.	1/1
37	Base arguments on concrete referents such as objects, drawings, diagrams, and actions.	2/2

Figure 30. Detail of Individual Student Report (ISR): Science Interim with Scoring Assertions

Marquez, Santos — **SCI Interim Grade 8 - Waves 2018-2019**

Student ID: 991007093 | Student DOB: 2/3/2002 | Enrolled Grade: 8 — Demo District 9998

Date Taken: 10/18/2018 — Demo School 999801

Score: 1/22

Item #	Scoring Assertion	Outcome
1-1	The student correctly identifies the change in amplitude recorded in the simulation, providing some evidence of student ability to observe and summarize how waves change in different media.	✕
1-2	The student correctly identifies that the frequency does not change, providing some evidence of student ability to identify wave properties and how they change in different scenarios.	✕
1-3	The student identifies that frequency does not change, providing some evidence of student ability to identify how properties of the medium affect each wave characteristic.	✕
1-4	The student identified that amplitude increases (or is indeterminate if they did not record it decreasing in the simulation), providing some evidence of student ability to identify how properties of the medium affect each wave characteristic.	✕
1-5	The student indicates that the wavelength cannot be determined simply from density (or that it increases if they found it to decrease in the simulation), providing some evidence of student ability to interpret the data given and make inferences about the affect of media density on each wave characteristic.	✕
1-6	The student correctly calculates and records the amplitude of the wave through salt water (4) providing some evidence of student ability to make observations and use them in a model of sound waves through media with different properties.	✕
1-7	The student correctly calculates the wavelength of salt water (13-15), providing some evidence of student ability to make observations and use them in a model of sound waves through media with different properties.	✕
1-8	The student indicates that the wave speed increases as the density increases based on the observations (or if they recorded the wavelength in salt water as shorter, then decrease, or "cannot tell" if they recorded equal wavelengths), providing some evidence of student ability to use data to identify how wave properties change in each scenario.	✕
1-9	The student correctly calculates and records the the amplitude of the wave through water (5) providing some evidence of student ability to make observations and use them in a model of sound waves through media with different properties.	✕
1-10	The student correctly calculates the wavelength of water (11-13), providing some evidence of student ability to make observations and use them in a model of sound waves through media with different properties.	✕
1-11	The student correctly identifies that the change in wavelength that was recorded in the simulation, providing some evidence of student ability to observe and summarize how waves change in different media.	✕
1-12	The student correctly identifies that the change in wave speed (which goes up with the wavelength recorded in the simulation), providing some evidence of student ability to observe and summarize how waves change in different media.	✕
1-13	The student indicated that the wave speed cannot be determined (or increase, if they found it to decrease in the simulation), providing some evidence of student ability to interpret the data given and make inferences about the effect of media density on each wave characteristic.	✓

Figure 31. Detail of Individual Student Report (ISR): Math Interim (ICA) with Longitudinal Graph

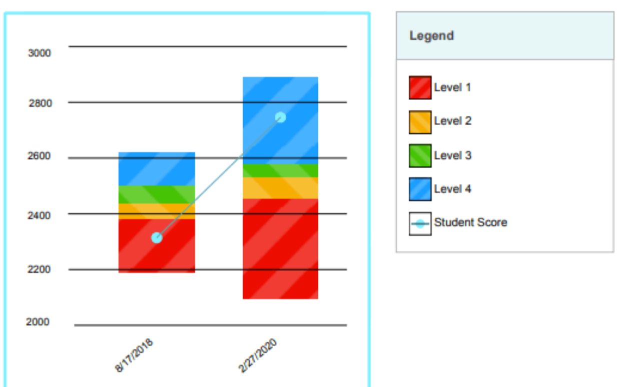
Jackson, Cynthia — **Grade 5 Mathematics - Interim (ICA) - Combined 2019-2020**

Student Name: Jackson, Cynthia | Student ID: 99996727 | Enrolled Grade: 5 — TestDISTRICT12

Date Taken: 2/27/2020 — TestSCHOOL9

Achievement Level: Level 4 **Scale Score:** 2746±58

Trend Report



Date	Test Reason	Test Label	Scale Score	Performance Level
8/17/2018 12:00:00 AM	Unassigned	Grade 3 MATH -COMBINED	2314±6	Level 1
2/27/2020 12:00:00 AM	Unassigned	Grade 5 Mathematics - Interim (ICA) - Combined	2746±58	Level 4

How to Generate and Export Individual Student Reports (ISR)

To generate and export an Individual Student Report (ISR) that details a student's performance on a test opportunity, use the Student Results Generator. You can select any combination of a test reason, assessments within a subject, and students in order to generate either a single ISR or multiple ISRs at once.

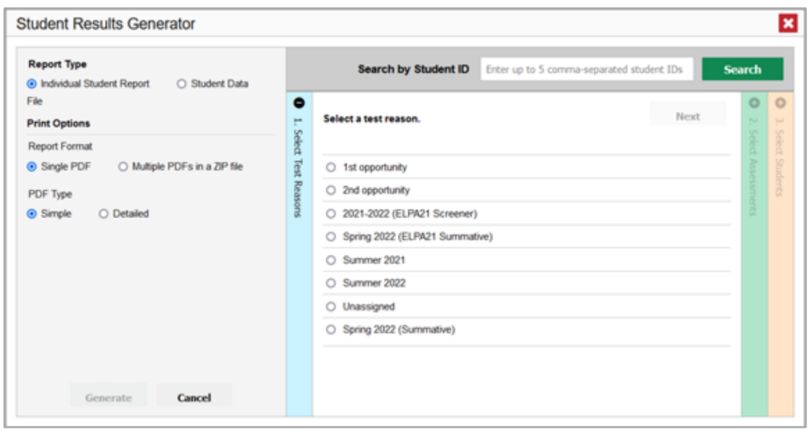
You may want to use the Student Results Generator to simultaneously print large numbers of ISRs.

ISRs can be generated from almost any Reporting page.

1. Click the **Download Student Results** button in the **Features & Tools** menu. The ***Student Results Generator*** window opens (Figure 32).

   Depending what page you open the Student Results Generator from, the options available to you may be prepopulated or preselected; for example, the Student Portfolio Report prepopulates a single student, and the Student Performance on Test report preselects a single test opportunity. (The filters applied to the page have no effect, however.) You can change the selections.

   Figure 32. Student Results Generator Window as Opened from Performance on Tests Report



2. In the panel on the left, select **Individual Student Report**. Always do this before you make other selections. Switching between the **Individual Student Report** and **Student Data File** options may revert some selections.

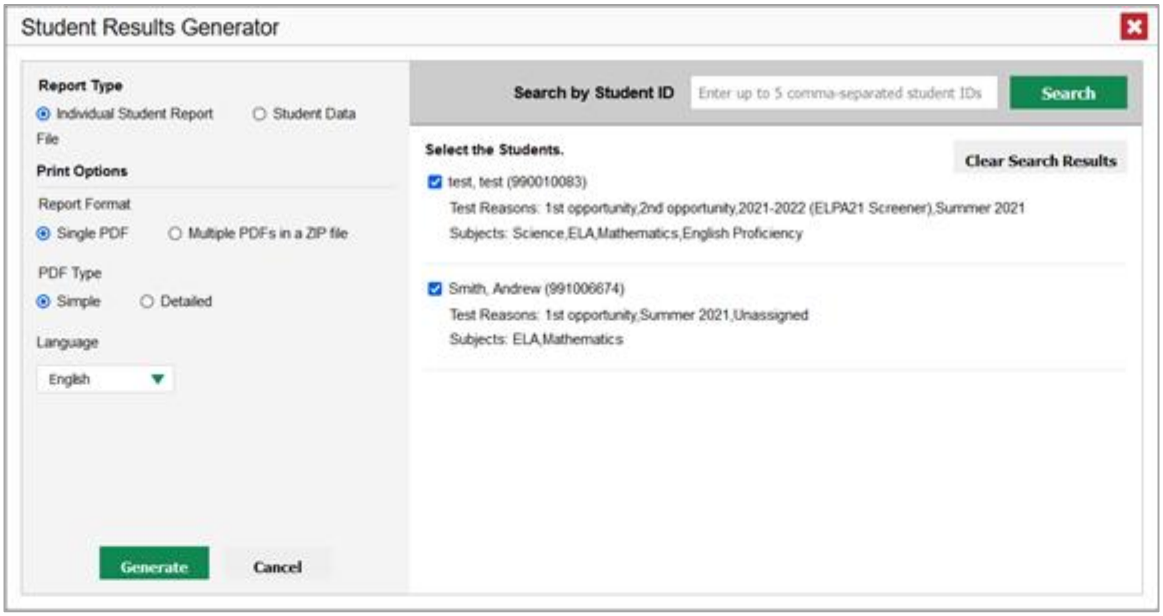
3. If you're generating multiple ISRs, then under Report Format, choose either a single PDF for all the ISRs, or a ZIP file containing a separate PDF for each one. If you select **Single PDF**, the Student Results Generator may nonetheless create a ZIP file of multiple PDFs depending on the number of schools, grades, and opportunities included.

4. Under PDF Type, select either a simple or detailed PDF.

5. If the test opportunity options are not preselected, or if you want to change them, there are two ways to make selections:

- Search for students. In the search field at the upper-right corner, enter up to 5 comma-separated student IDs and click **Search**. The resulting list of students and all the tests they've taken will replace any previous selections, as in Figure 33. To deselect and clear results, click **Clear Search Results**.

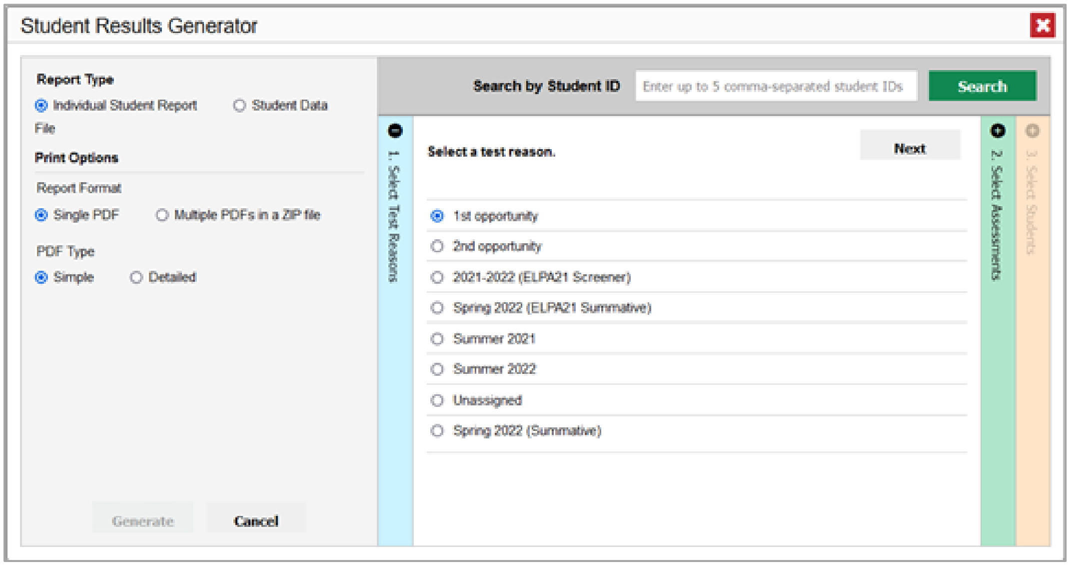
Figure 33. Student Results Generator Window: Student Search Results



- Use the three accordion sections. (If a student is prepopulated by the Student Portfolio Report, first remove the student by clicking **Clear Search Results**.) Starting from the left, click the section bars to expand them or use the **Next** and **Previous** buttons to navigate. Within each section you must make selections using the radio buttons and checkboxes:

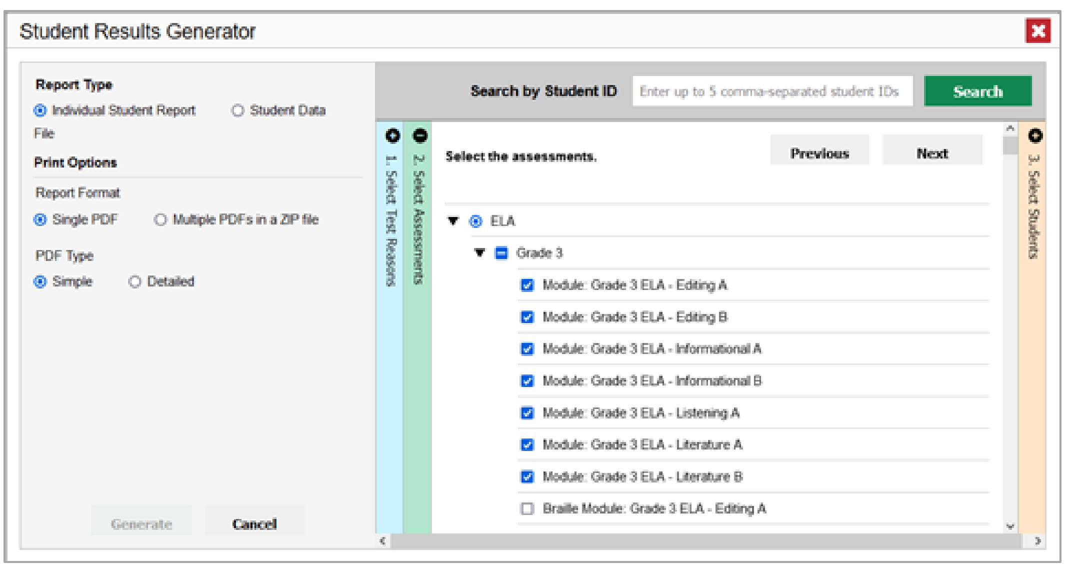
i. In the **Select Test Reason** accordion section (Figure 34), choose a test reason. Test reasons are either test windows or categories for tests.

Figure 34. Student Results Generator Window: Select Test Reason Section



ii. In the **Select Assessments** section (Figure 35), choose any number of tests or grade levels within a single subject.

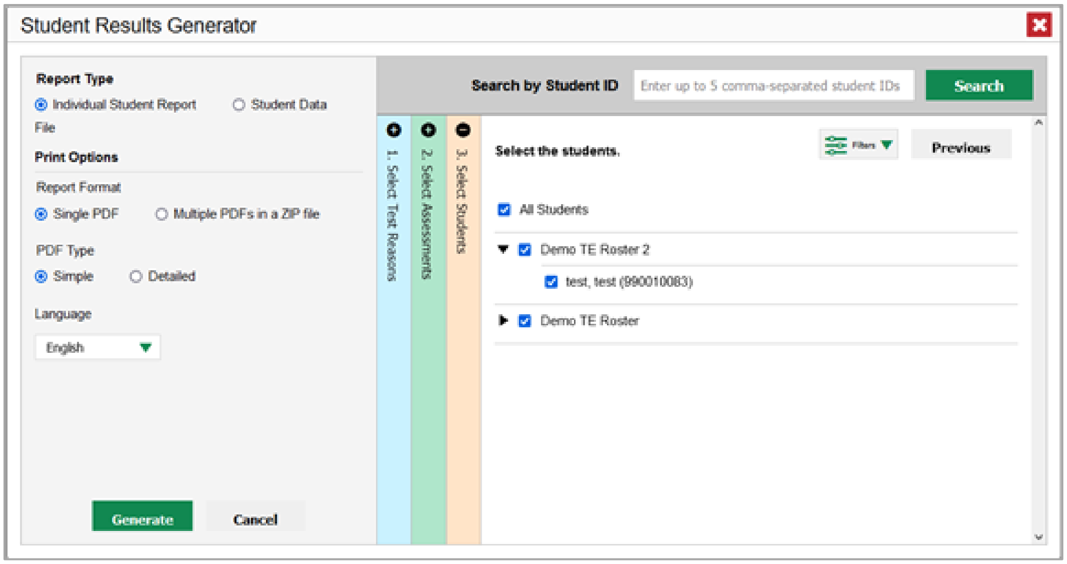
Figure 35. Student Results Generator Window: Select Assessments Section



iii. In the **Select Students** section (Figure 36), choose students from the expandable school and/or class (roster) options. If you're a district-level user, you may choose students in up to three schools.

- Sometimes the list of students is truncated. You can display the entire list by clicking **Click to Load More**.
- Note that marking the checkbox for a student in one class (roster) or school also marks it anywhere else the student appears, and the same goes for clearing the checkbox.

Figure 36. Teacher View: Student Results Generator Window: Select Students Section



6. *Optional*: If you used the accordion sections to make your selections, then to set a range of processing dates for which to generate results, use the filter menu as follows:

    a. Open the **Filters** menu. The menu displays two date fields, as in Figure 37.

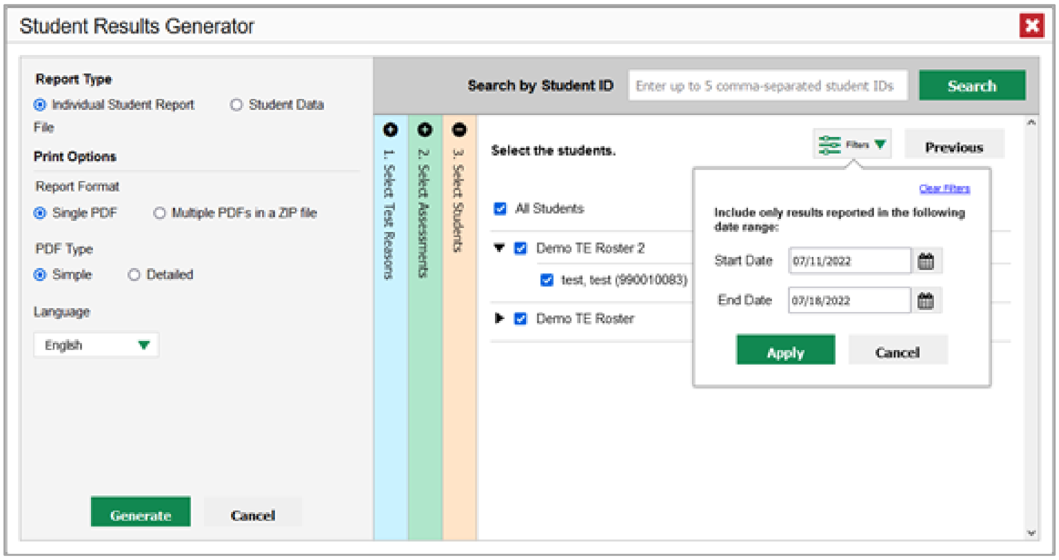
    b. Use the calendar tools to select dates, or enter them in the format `mm/dd/yyyy`.

    c. Click **Apply**.

    d. *Optional*: To revert to including results for all available dates, reopen the filter menu, click **Clear Filters**, then click **Apply**.

Note that processing date is not always the same as the date a test was taken.

Figure 37. Teacher View: Student Results Generator Window: Select Students Section with Filters Menu Open



7. Click **Generate**. Once ISR generation is finished, the Inbox contains the new ISR(s) available for download.

Note that if a student took a test multiple times with different test reasons, an ISR will be generated for each test opportunity. If a student took a test multiple times with the same test reason, only the most recent test opportunity will be included. You can create an ISR for another test opportunity by navigating directly to the report for that opportunity. Past test opportunities are marked with numbers ① in reports, starting with the earliest.

How to Generate and Export Student Data Files

This section discusses student data files, which are useful for analysis.

To generate and export student data files, use the Student Results Generator. You can select any combination of a test reason, assessments, and students in order to generate and export the files.

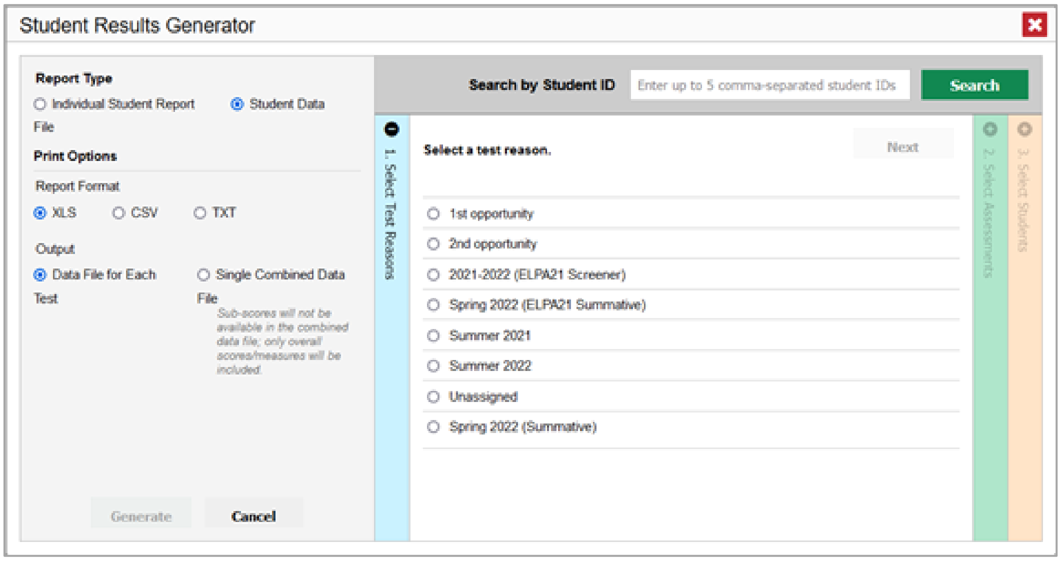
You can generate student data files from almost any report page.

1. Click the **Download Student Results** button in the **Features & Tools** menu. The ***Student Results Generator*** window opens.

   Depending what page you open the Student Results Generator from, the options available to you may be prepopulated or preselected; for example, the Student Portfolio Report prepopulates a single student, and the Student Performance on Test report preselects a single test opportunity. (The filters applied to the page have no effect, however.) You can change the selections.

2. In the panel on the left, select **Student Data File**, as in Figure 38. Always do this before you make other selections. Switching between the **Individual Student Report** and **Student Data File** options may revert some selections.

Figure 38. Student Results Generator Window



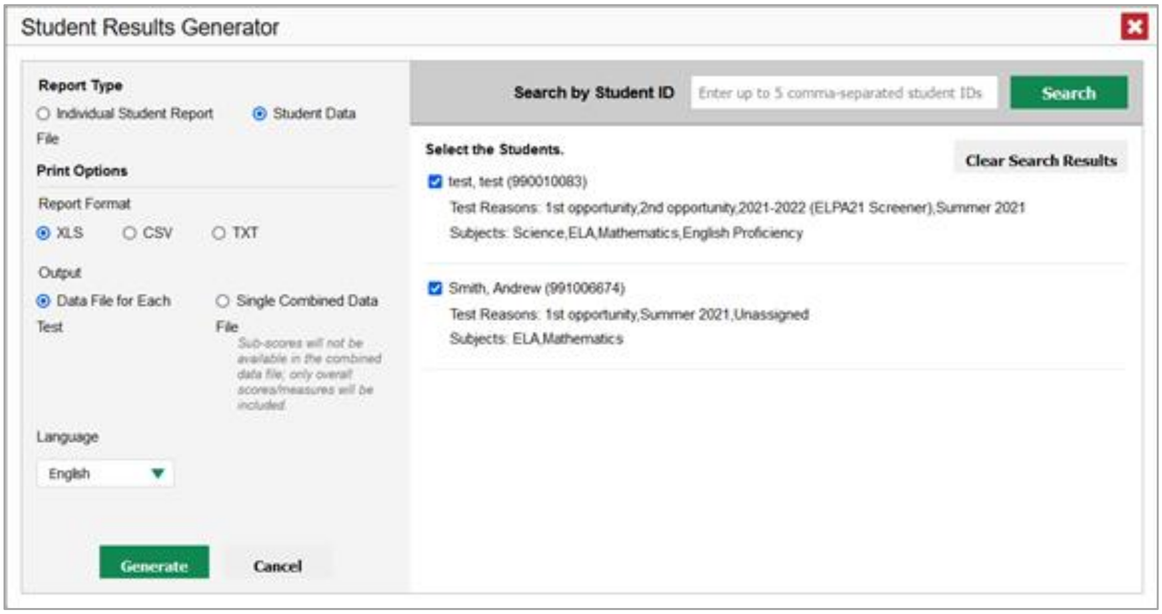
3. Under Report Format, select **XLS** (Excel .xlsx), **CSV** (comma-separated values), or **TXT** (tab-delimited text).

4. Under Output, select either **Data File for Each Test** or **Single Combined Data File**. Note that a single combined file does not include reporting categories. Large files may be split by school.

5. If the test opportunity options are not preselected, or if you want to change them, there are two ways to make selections:

   - Search for students. In the search field at the upper-right corner, enter up to 5 comma-separated student IDs and click **Search**. The resulting list of students and all the tests they've taken will replace any previous selections, as in Figure 39. To deselect and clear results, click **Clear Search Results**.

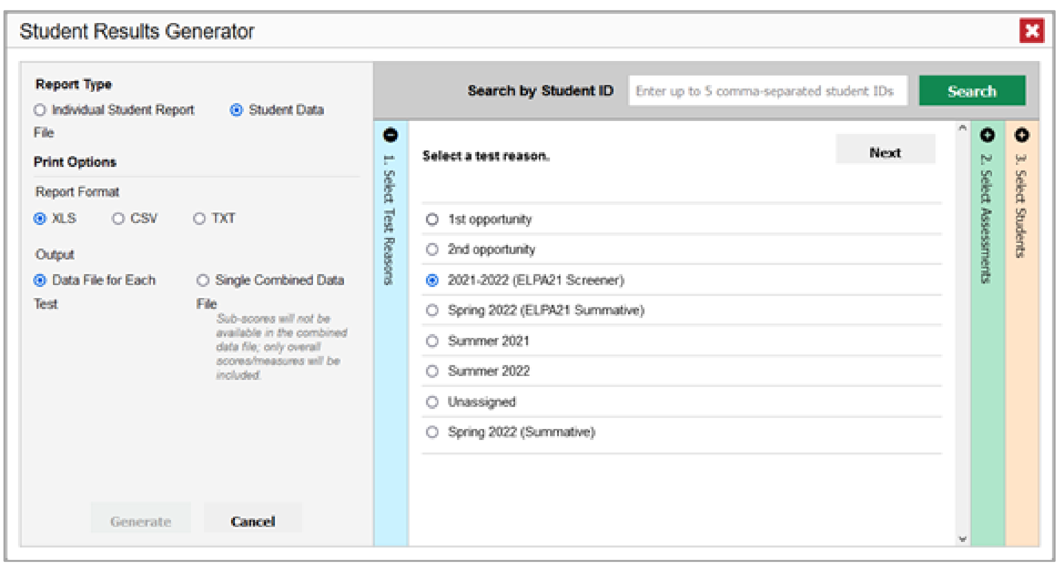
Figure 39. Student Results Generator Window: Student Search Results



- Use the three accordion sections. (If a student is prepopulated by the Student Portfolio Report, first remove the student by clicking **Clear Search Results**.) Starting from the left, click the section bars to expand them or use the **Next** and **Previous** buttons to navigate. Within each section you must make selections using the radio buttons and checkboxes:

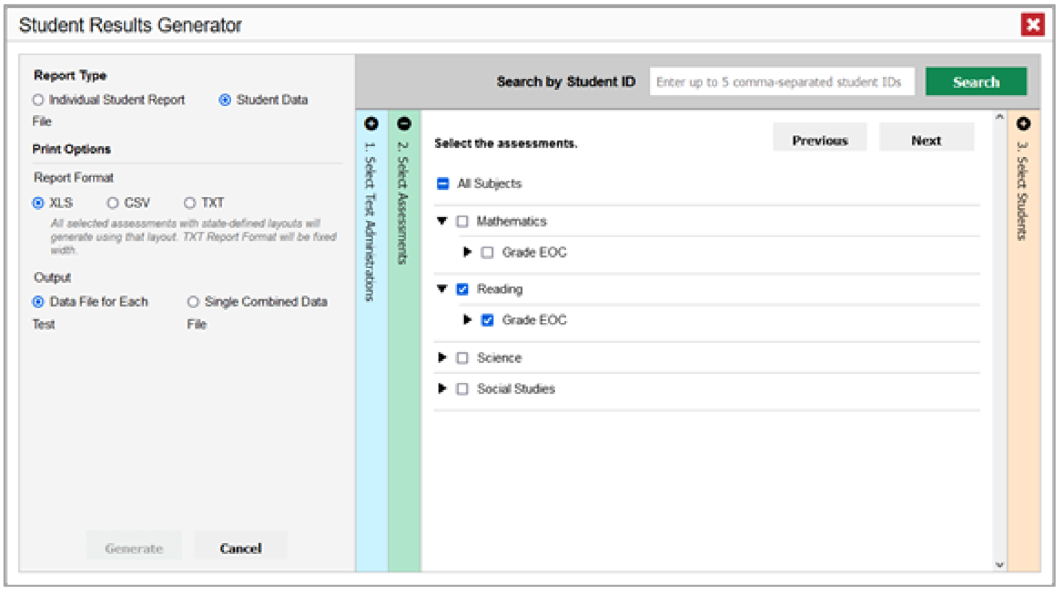
    i. In the **Select Test Reason** section (Figure 40), choose a test reason. Test reasons are either test windows or categories for tests.

Figure 40. Student Results Generator Window: Select Test Reason Section



    ii. The **Select Assessments** section (Figure 41) groups tests by subject and grade. Mark the checkboxes beside the tests or groups of tests you want to include in the report, or mark **All Subjects**.

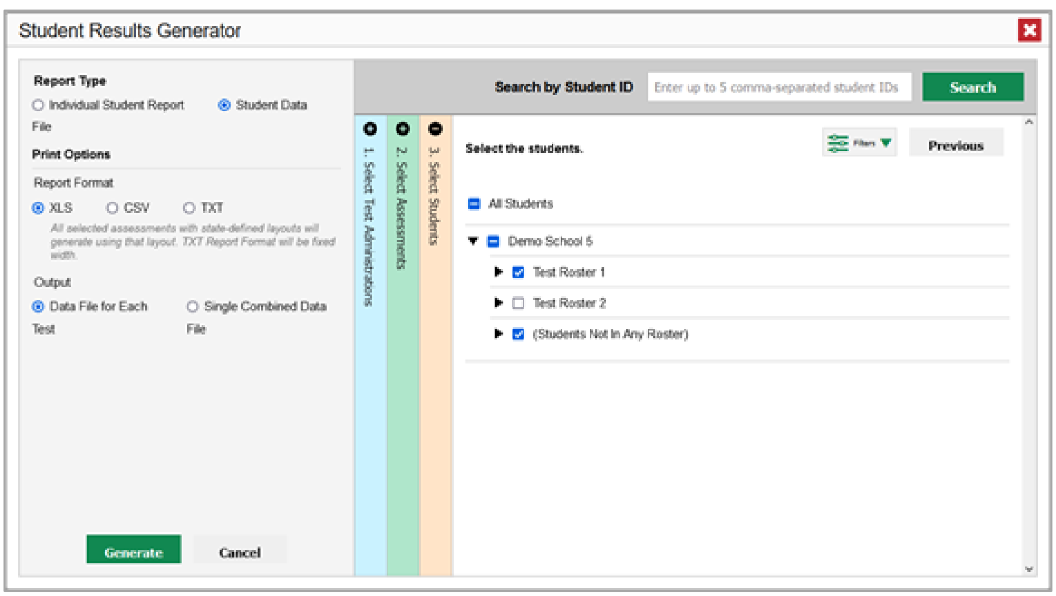
Figure 41. Student Results Generator Window: Select Assessments Section



iii. In the **Select Students** section ([Figure 42](#)), select any number of students from the expandable school and/or class (roster) options.

- Sometimes a list of students is truncated. You can display the entire list by clicking **Click to Load More**.
- Marking the checkbox for a student in one class (roster) or school also marks it anywhere else the student appears, and the same goes for clearing the checkbox.

Figure 42. District-Level User View: Student Results Generator Window: Select Students Section



7. *Optional*: If you used the accordion sections to make your selections, then to set a range of processing dates for which to generate results, use the filter menu as follows:

a. Open the **Filters** menu (see Figure 43). The menu displays two date fields.

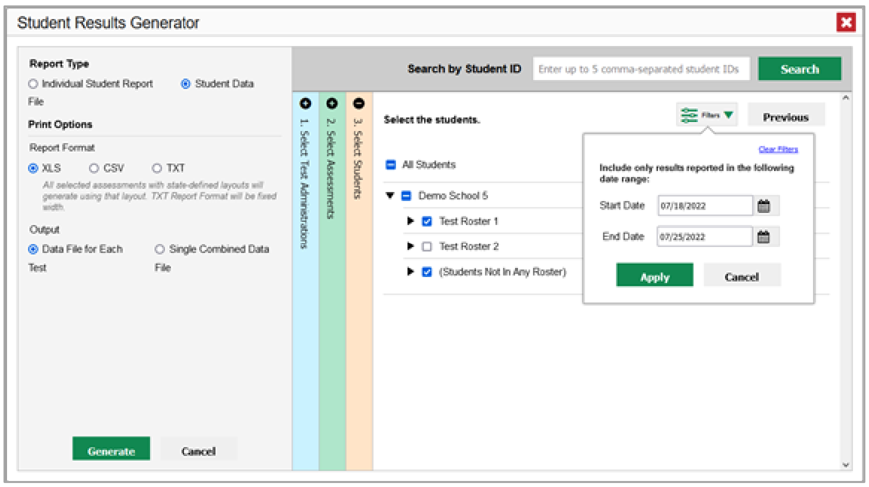
b. Use the calendar tools to select dates, or enter them in the format `mm/dd/yyyy`.

c. Click **Apply**.

d. *Optional*: To revert to including results for all available dates, reopen the filter menu, click **Clear Filters**, then click **Apply**.

Note that processing date is not always the same as the date a test was taken.

Figure 43. District-Level User View: Student Results Generator Window: Select Students Section with Filters Menu Open



8. Click **Generate**. Once data file generation is finished, the Inbox contains the new student data file(s) available for download.

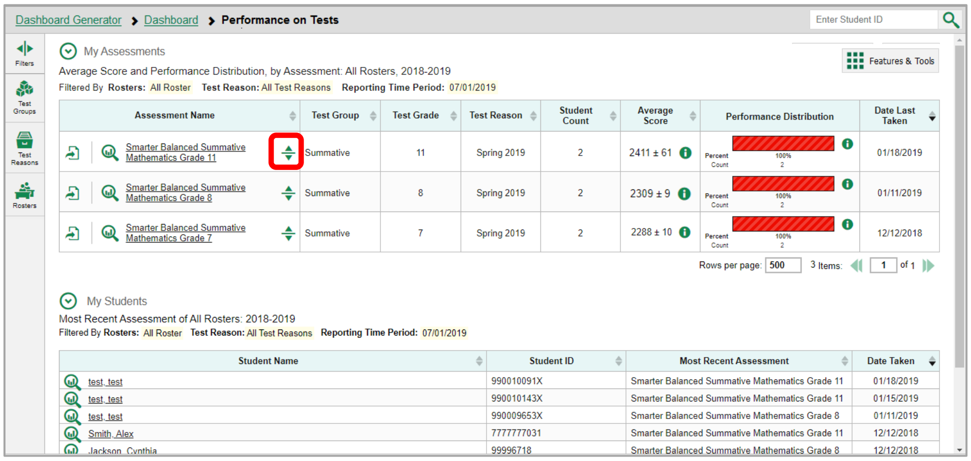
Note that if a student took a test multiple times, the files will include each test opportunity.

How to Compare Students' Data with Data for Your State, District, School, and/or Total Students

In the Performance on Tests report and the Student Portfolio Report, you can access performance data for your state, district, school, and/or total students.

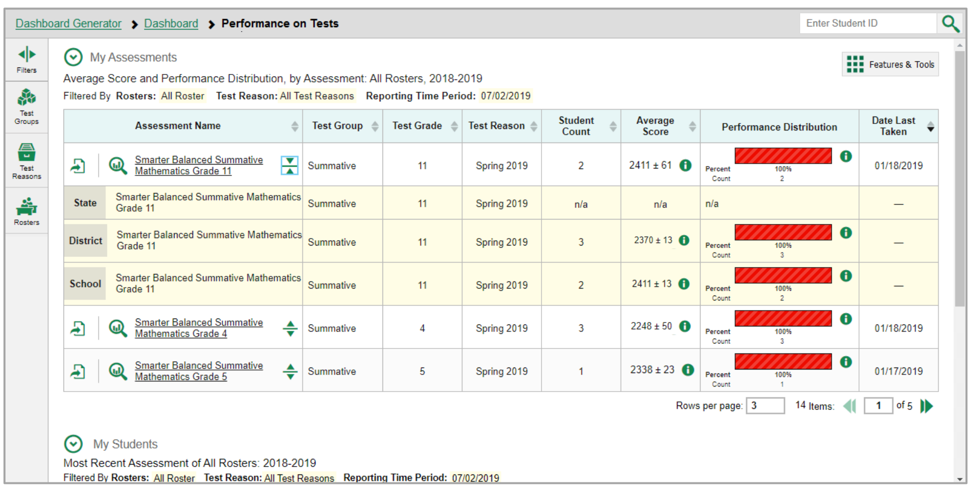
In the Performance on Tests report (see Figure 44), click to the right of a test name.

Figure 44. Teacher View: Performance on Tests Report



Rows containing data for the state, district, and/or school appear below, as in Figure 45.

Figure 45. Teacher View: Performance on Tests Report with Expanded Comparison Rows



To hide the comparison rows, click to the right of the test name.

The comparison feature also appears on the Student Portfolio Report page, to generate comparisons for a student's performance on any test with that of your state, district, school, and/or total students.

How to Set Up Reports for Summatives and Interims to Suit Your Needs

You can set up your reports so it’s easier to access the data that are most important to you. For example, if you’re a teacher, you may want to hide certain tests in subjects you don’t teach, or you may want to narrow down your reports to a single roster.

This section explains how to make several different adjustments to reports: showing only the tests you’re interested in; showing only the classes (rosters) you’re interested in; showing only the schools you’re interested in; and viewing data from a previous point in time.

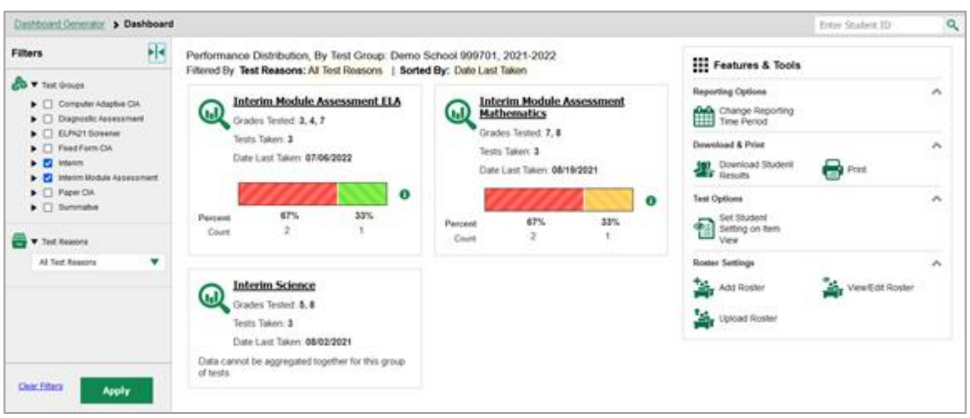
How to Filter Tests to Display

You can filter the tests you want to view in reports. You may want to do this, for example, if you are an ELA teacher and you do not want to see your students’ math scores. By default, the data for those mathematics assessments appear in your reports.

Filtering tests to display begins on the Dashboard Generator page. This is where you can select the test groups you want to view on the dashboard and, as an option, set those selections as defaults. You can also temporarily filter the tests that appear in the reports you are currently viewing, as described below.

1. On the left side of the dashboard or the Performance on Tests report, click either the **Filters** panel expand button or the **Test Groups** button. The **Filters** panel expands (see Figure 46).
2. Mark as many selections as you like in the **Test Groups** section of the filters panel (see Figure 46). Tests are organized by test type, subject, and grade.

Figure 46. Teacher View: Dashboard with Expanded Filters Panel



3. Click **Apply**. The report updates to show only data for those tests.
4. *Optional*: To revert all filters to their defaults, open the filters panel again and click **Clear Filters**. Click **Apply**. Filters also revert when you log out, switch user roles, or switch systems.

For Teachers and School-Level Users: How to Filter Classes (Rosters) to Display

Rosters usually represent classes, but can represent any group that's meaningful to users, such as students who have taken honors courses. Each roster is associated with a teacher. To learn how to create and modify rosters in the Reporting System, see Class (Roster) Management.

In the Performance on Tests report, teachers and school-level users can filter by a particular roster. When you filter, you eliminate students not in the selected class from the data you're viewing.

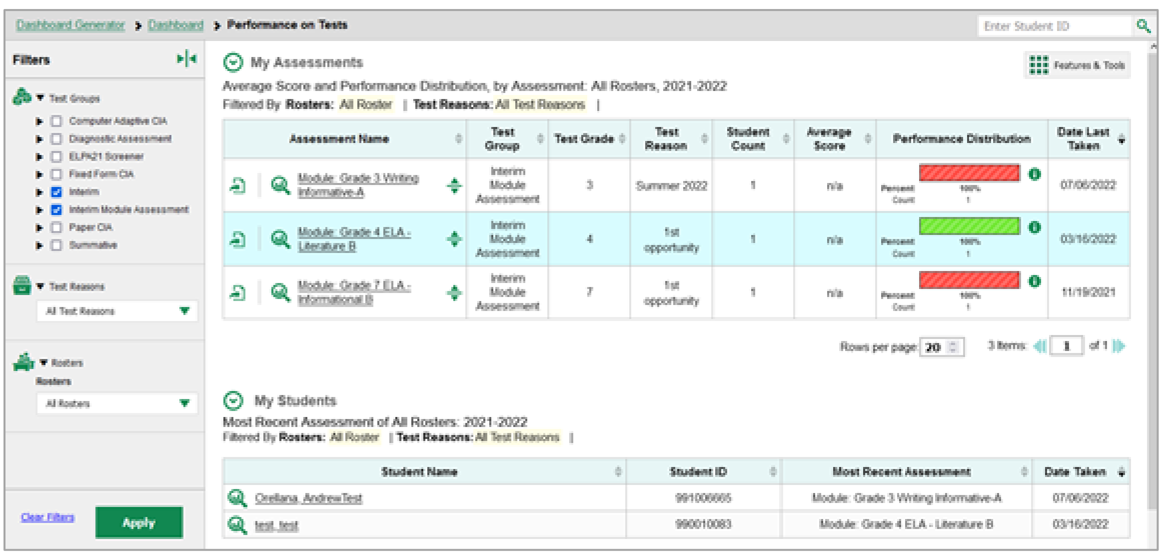
Filtering by roster makes it easy to focus on a particular class's performance. And by switching filters, you can easily compare one class with another. If you don't filter by roster, the reports default to showing data for all classes. You may find data for a single class easier to understand.

1. On the left side of the Performance on Tests report, click either the **Filters** panel expand button or the **Rosters** button. The **Filters** panel expands (see Figure 47).

2. Make a selection from the drop-down list in the **Rosters** section.

   - If you're a school-level user, you must first select a teacher from the drop-down list, and then select a particular class (roster) from the second drop-down list that appears. By default the first class listed is selected.

Figure 47. Teacher View: Performance on Tests Report with Expanded Filters Panel



3. Click **Apply**. The report updates to show only data for that class (roster).

4. *Optional*: To revert all filters to their defaults, open the filters panel again and click **Clear Filters**. Click **Apply**. Filters also revert when you log out, switch user roles, or switch systems.

All the reports accessible from this page will be filtered the same way.

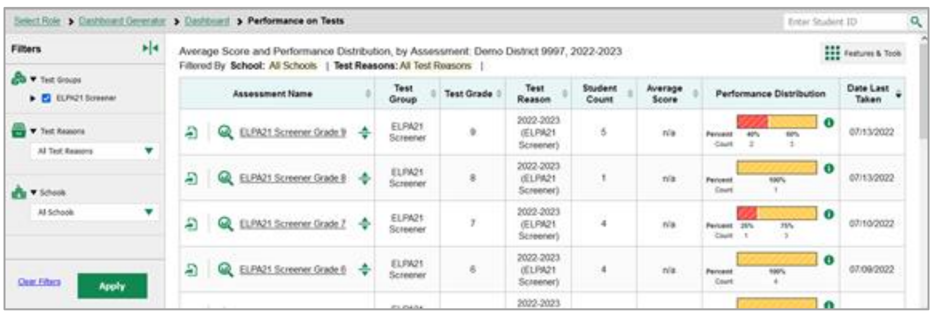
The row of filter details below each table header shows the classes (rosters) you're viewing.

For District-Level Users: How to Filter Schools to Display

Filtering the Performance on Tests report by school makes it easy to focus on a particular school's performance. And by switching filters, you can easily compare it with another school. If you don't filter by school, the Performance on Tests report defaults to showing data for all schools. You may find data for a single school easier to understand.

1. On the left side of the Performance on Tests report, click either the expand button or the **Schools** button . The **Filters** panel expands (see Figure 48).
2. Make a selection from the drop-down list in the **Schools** section (see Figure 48).

Figure 48. District-Level User View: Performance on Tests Report with Expanded Filters Panel



3. Click **Apply**. The report updates to show only data for that school.
4. *Optional*: To revert all filters to their defaults, open the filters panel again and click **Clear Filters**. Click **Apply**. Filters also revert when you log out, switch user roles, or switch systems.

If you click the name of a test (or beside it) when you've filtered by a single school, the link will take you to the School Performance on Test report and not to the District Performance on Test report.

The row of filter details below the table header shows the schools you're viewing.

How to View Data from a Previous Point in Time

Changing the reporting time period allows you to view test results from a previous point in time. There are two time period settings: you can select a school year for which to view tests, and you can enter a date for which to view students.

- When you set a school year for which to view tests, the reports show data for test opportunities completed *in the selected school year*.
- When you set a date for which to view students, the reports show data only for the students who were associated with you *as of the selected date*. Students' enrollment and demographic

information is all given as of the selected date as well. You can use this setting to view data for students who have left or recently entered your classes (rosters), school, or district.

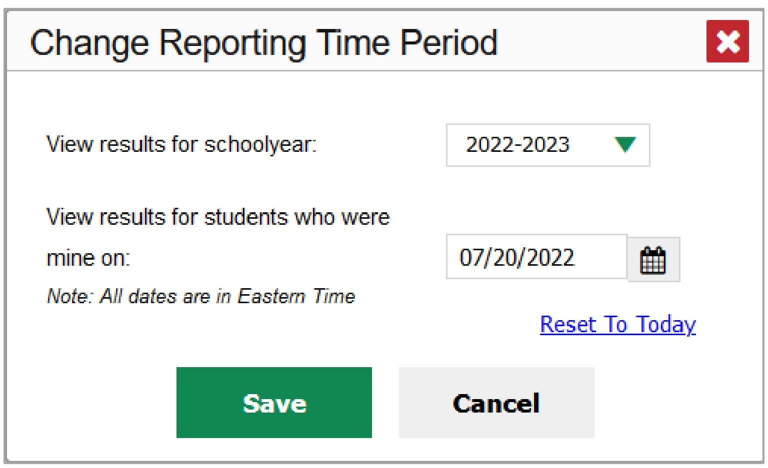
If you don't change the reporting time period, or if you reset it to the default, all the reports show test opportunities only for the current school year (except Longitudinal Reports and Student Portfolio Reports, which always retain the ability to look back to previous years), with current student data.

Some examples of how you can use this feature:

- You may want to view the past performance of your current students, including new transfer students. In that case, set a school year in the past and keep the date set to today.
- You may want to view the performance of your former students in order to compare them with that of your current students. In that case, set the date to a time when your former students belonged to you and had started testing, and set the school year to the same time. Then switch back to the present to compare.

1. From the **Features & Tools** menu, select **Change Reporting Time Period**. (If you're viewing the Dashboard Generator page, click **Change the reporting time period.** The dashboard also offers this link when no assessments are available to display.) The ***Change Reporting Time Period*** window appears (see Figure 49).

Figure 49. Change Reporting Time Period Window



2. From the school year drop-down list, select a school year (see Figure 49). This is the year for which you will view test results.

3. In the *View results for students who were mine on* field, use the calendar tool to select a date, or enter it in the format `mm/dd/yyyy`. You will be viewing all the students who were associated with you on that date, and only those students.

   - To view your current students' past performance, keep the date set to today.
   - To view the performance of your former students, set the date to a day when those students were associated with you and had started testing.

4. Click **Save**. All reports are now filtered to show only data for the selected school year and date. The selected date displays in the filter details below the report headings. All other filters are cleared.

5. *Optional*: To go back to viewing the latest data, open the ***Change Reporting Time Period*** window again, click **Reset To Today** in the lower-right corner, then click **Save**. The date resets and all filters are cleared. The reporting time period also resets when you log out, but persists when you switch roles.

How to Export and Print Data on Summatives

You can export or print any data you see in the Reporting System. Some can be exported directly from the Performance on Tests report. You may want to export or print to save a snapshot of data to consult later, or to share data. Different options will be available depending on the report you are viewing.

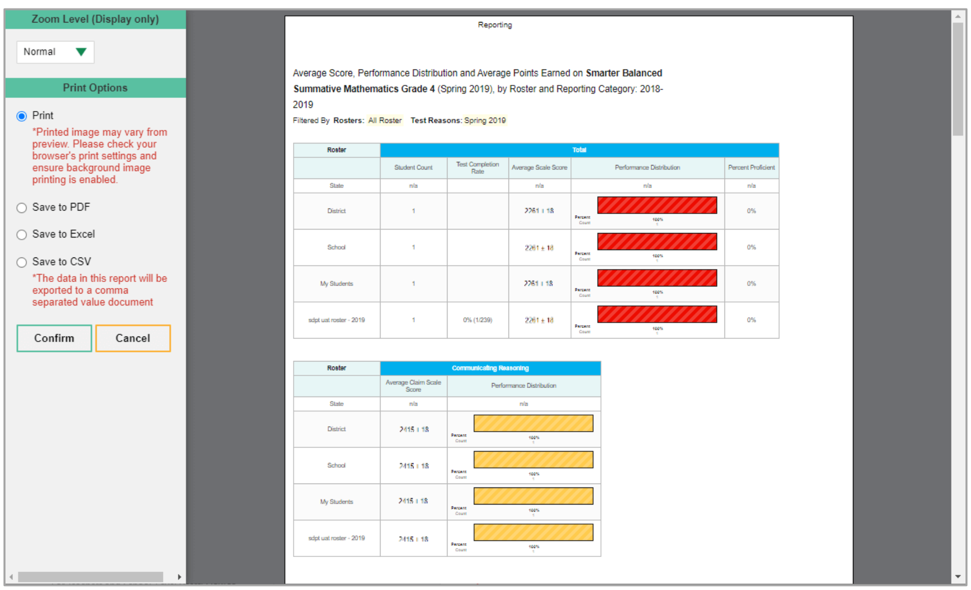
How to Export or Print a Report You're Viewing

1. Select the print button from the **Features & Tools menu** or, in a Longitudinal Report window, from the upper-right corner. If there are multiple report tables on the page, multiple print buttons appear.

A print preview page opens (see Figure 50).

- To zoom in on the print preview, use the drop-down list under the *Zoom Level (Display only)* section. This setting affects the preview only.

Figure 50. Print Preview Page



2. Do one of the following under the *Print Options* section:

- To print the report, select the **Print** radio button.
- To download a PDF version of the report, select **Save to PDF**. Then select an option from the **Page Layout** drop-down list that appears.
- To download a Microsoft Excel (.xlsx) version, select **Save to Excel**.
- To download a comma-separated value (CSV) version of the report, select **Save to CSV**.

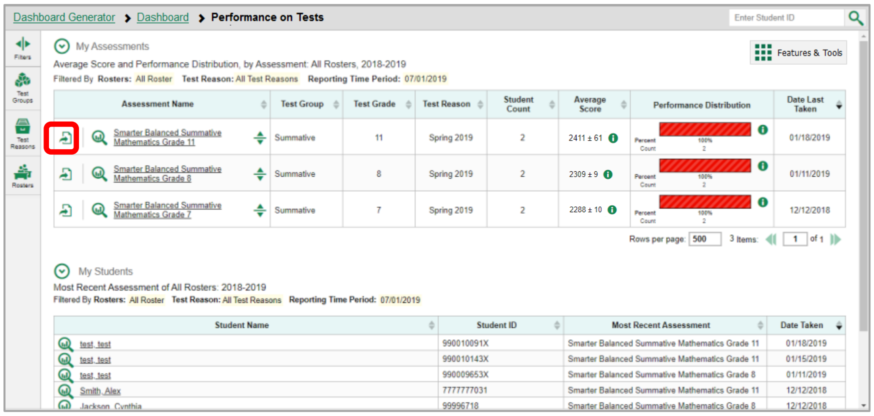
3. Click **Confirm**.

If you saved the report as PDF, Excel, or CSV, the ***Secure Inbox*** window appears, displaying the generated report.

How to Export an Assessment Report Directly from the Performance on Tests Report

1. Click the export button to the left of the name of the assessment whose report you wish to export (see Figure 51).

Figure 51. Teacher View: Performance on Tests Report

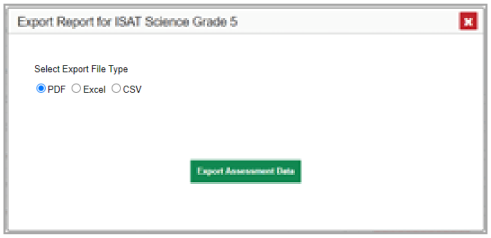


The ***Export Report*** window opens. The options in this window vary according to your user role (see Figure 52 and Figure 53).

2. If necessary, select which report to export for the assessment.

   - **Teachers and school-level users**: The exported report will contain test results for all your students.

Figure 52. Teacher View: Export Report Window

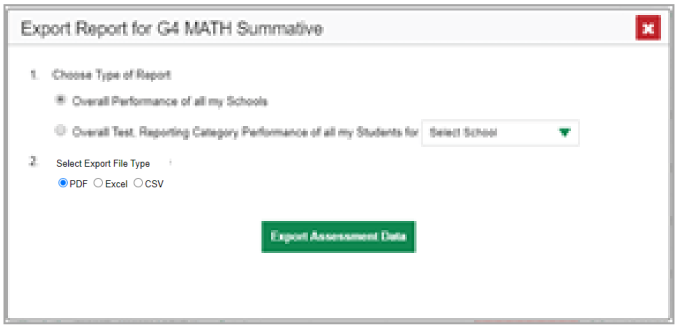


   - **District-level users**: Select which report to export for the assessment (see Figure 53).

     – To export the district test results for the assessment, mark the **Overall Performance of all my Schools** radio button.

     – To export school test results, mark the **Overall Test & Reporting Category Performance of all my Students for [School Name]** radio button, then select a school from the drop-down list.

Figure 53. District-Level User View: Export Report Window



5. Do one of the following:

- To export the report in PDF format, mark the **PDF** radio button.
- To export the report in .xlsx format, mark the **Excel** radio button.
- To export the report in comma-separated values (CSV) format, mark the **CSV** radio button.

6. Click **Export Assessment Data**. A confirmation window appears.

7. Click **Yes** to export or **No** to return to the ***Export Report*** window. When you've exported a file, the ***Secure Inbox*** window appears with the generated file available for download.

More About How to Work with Interims

This section explains some Reporting System features and functions that are specific to interim assessment reports. These features cannot be used with summative assessment reports.

How to Access Item-Level Data on Interims

Unlike summatives, interim and benchmark assessments contain non-secure, non-public items. Reports for individual interim and benchmark tests include the following:

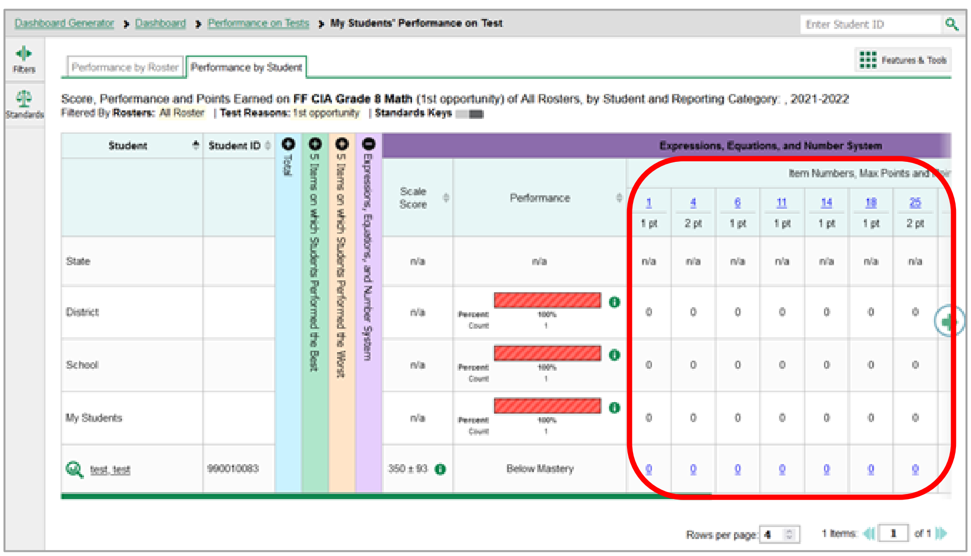
- Item-level data.
- Access to the items themselves.
- Access to student responses to the items.

Test results for adaptive assessments include item-level data only on the individual student level.

How to View Item Scores

To expand sections containing item data, click the vertical section bars as in Figure 54.

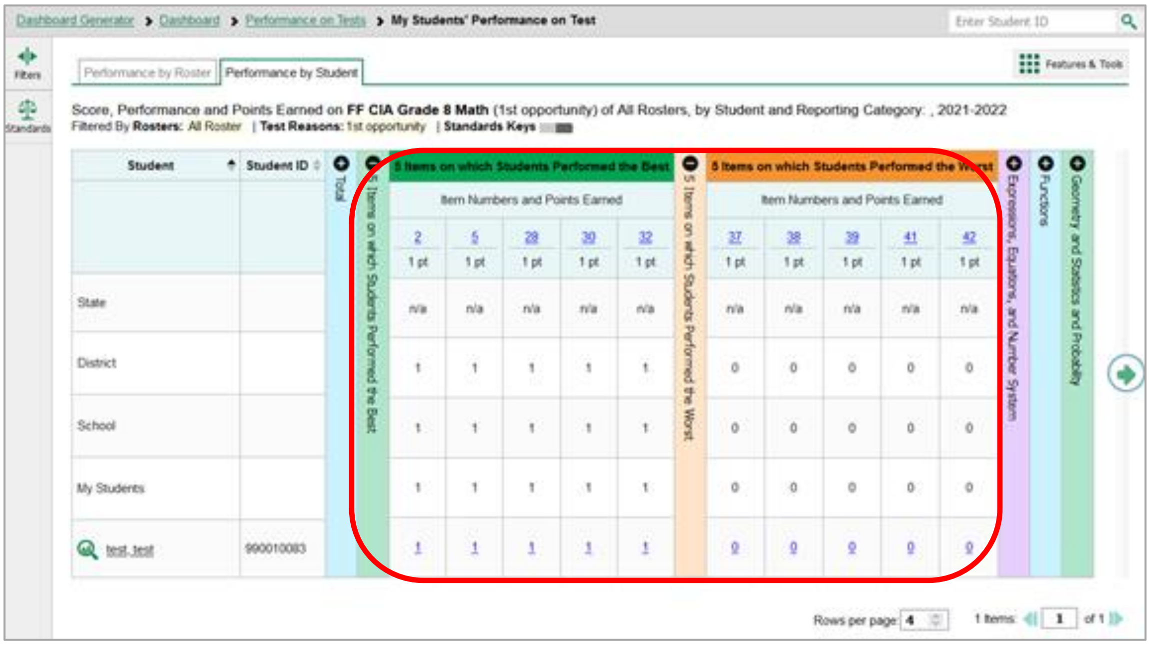
Figure 54. My Students' Performance on Test Report: Performance by Student Tab with Expanded Reporting Category Section



How to Find Out Which Items Students Performed on the Best or Struggled with the Most

Look in the sections **5 Items on Which Students Performed the Best** and **5 Items on Which Students Performed the Worst** (see Figure 55). You can click the vertical section bars to expand them, just like other sections.

Figure 55. My Students' Performance on Test Report: Performance by Student Tab with Expanded 5 Items on Which Students Performed the Best and Worst Sections

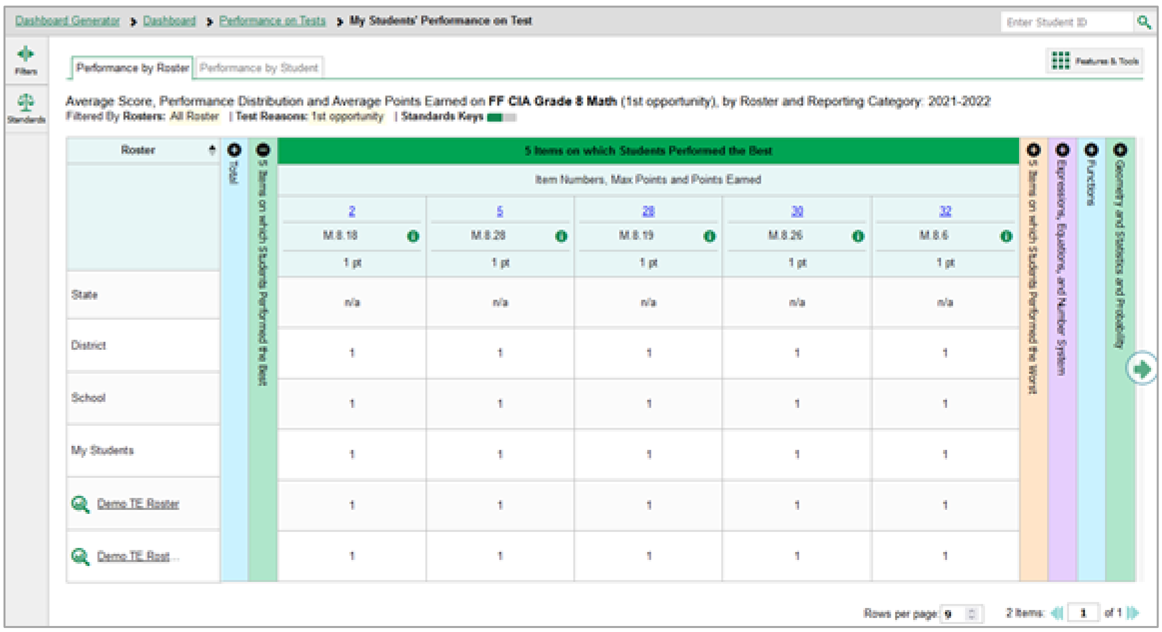


How to View Standards for Each Item

In a report displaying item-level data, you can view the standard or standards to which each item is aligned. This allows you to determine at a glance what the item measures.

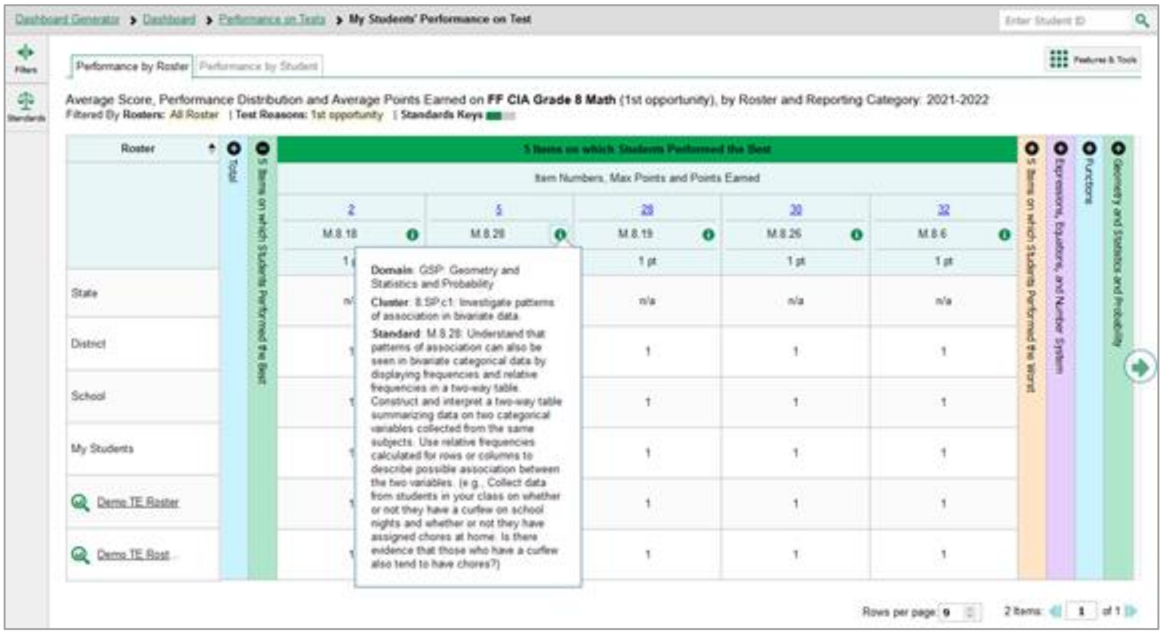
To show and hide item standards, click the **Standards Keys** toggle in the row of filter details below the report table heading. Under each item number appears a standard key or list of standard keys (see Figure 56). Note that this toggle does not affect printouts or exports, which always include the standard keys when they include item-level data.

Figure 56. My Students' Performance on Test Report with Expanded 5 Items on Which Students Performed the Best Section



Click the more information buttons beside the standard keys to view legends displaying the full text of each cluster (category of standards) and each standard, as in Figure 57. This full text is not included in printouts or exports.

Figure 57. My Students' Performance on Test Report with Expanded Reporting Category Section and Expanded Legend



How to View an Item

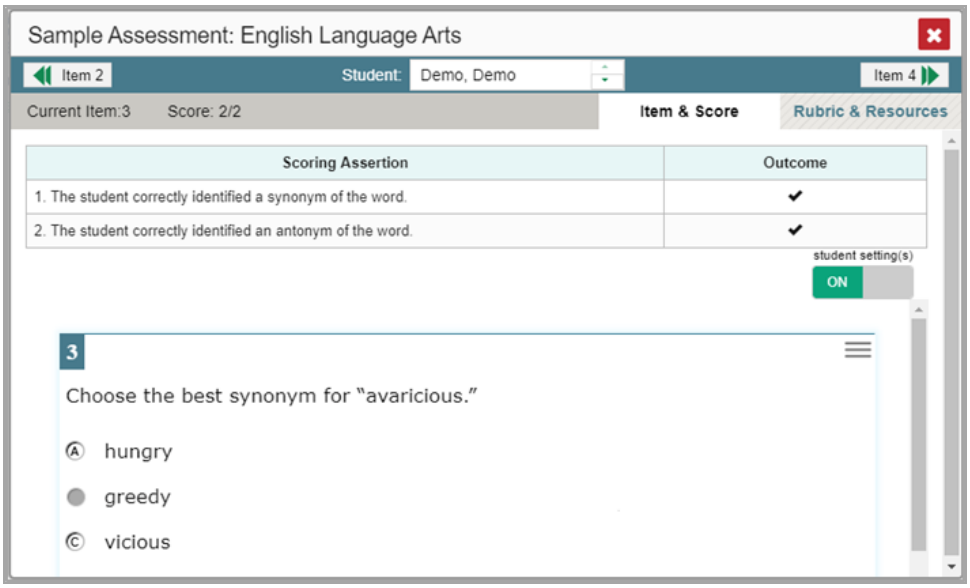
You can view the actual items themselves, along with student responses to those items.

Do either of the following:

- To view the item in a blank state, click the item number in the first row of the report table.
- To view the student's response to the item, find that student's name in the Student column on the left. Then click the score the student obtained on that item.

The ***Item View*** window appears (see Figure 58). It contains an **Item & Score** tab and a **Rubric & Resources** tab. A banner at the top of the window displays the item's number, score (when the item includes the student's response), and confidence level (when a machine-suggested score has a low confidence level). The **Item & Score** tab (see Figure 58) shows the item and may include a particular student's response.

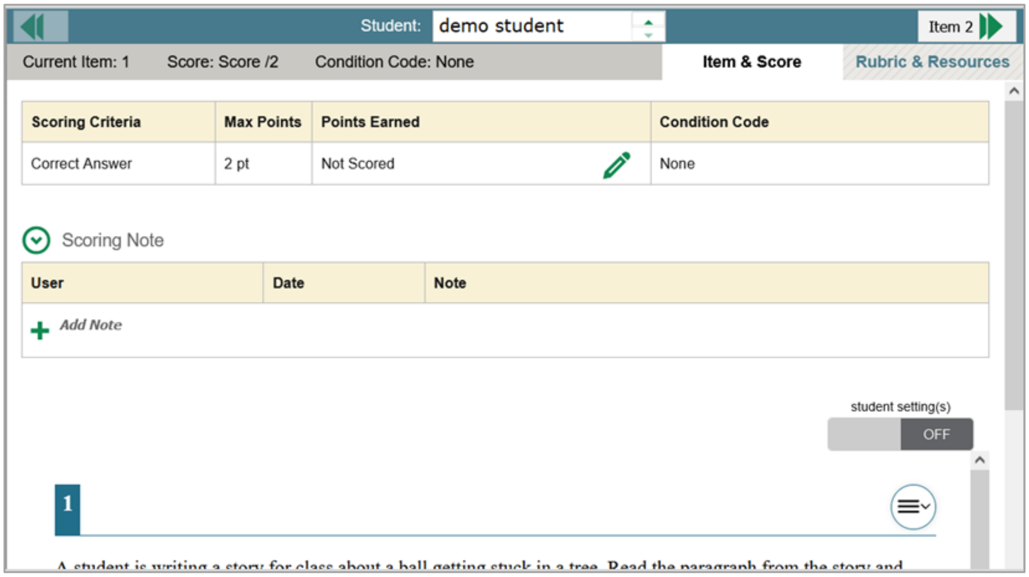
Figure 58. Item View Window: Item & Score Tab with Student Response



The **Item & Score** tab may include the following sections.

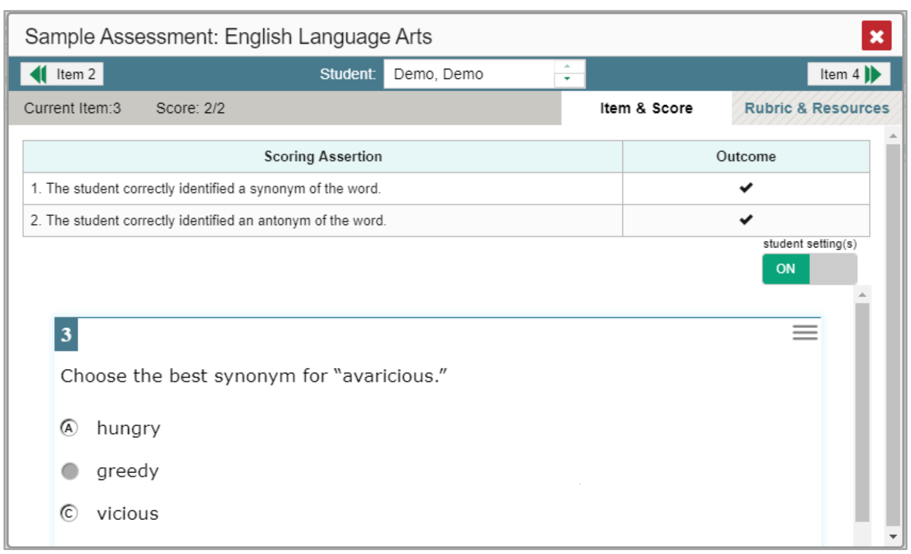
- **Scoring Criteria:** When you're viewing a student's response and the item has scoring criteria, the Scoring Criteria table (see Figure 59) lists the name, maximum points, points earned, and condition codes for each scoring criterion. This table also allows you to modify scores for items with editable scores. Note that for some items, a second Scoring Criteria table appears, displaying transformed scores.

Figure 59. Item View Window: Item & Score Tab with Student Response and Scoring Criteria Table



- **Scoring Assertion:** Each scoring assertion contains both a statement that provides information about what the student did in their response, and the content knowledge, skill, or ability that is evidenced by their response. When you're viewing a student's response and the item has scoring assertions, the Scoring Assertion table appears, listing each assertion and outcome (see Figure 60).

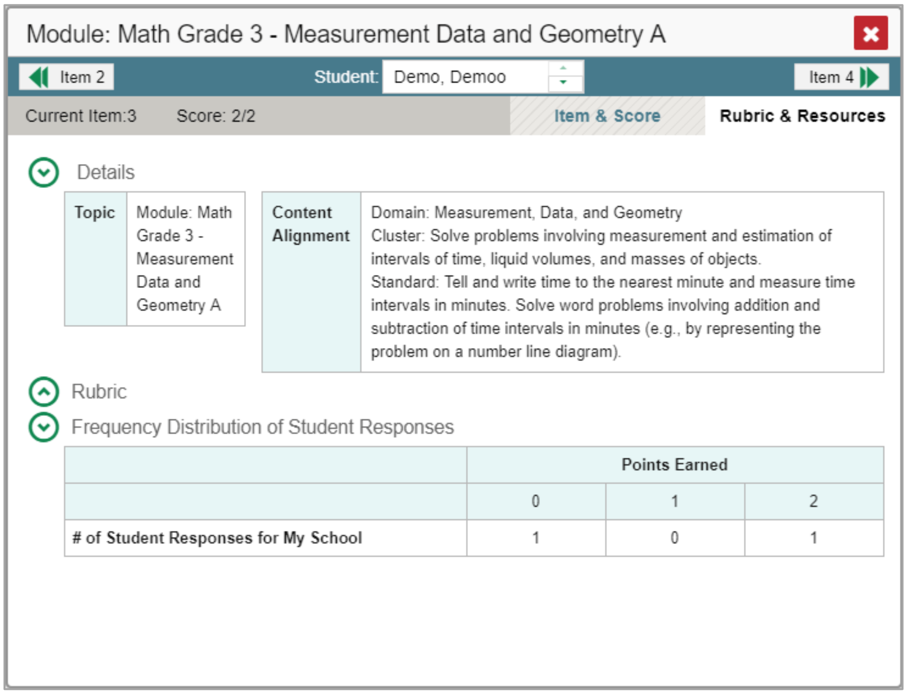
Figure 60. Item View Window: Item & Score Tab with Student Response and Scoring Assertion Table



- **Item:** Displays the item as it appeared on the assessment in the Student Testing Site. For items associated with a passage, the passage also appears.

The **Rubric & Resources** tab (see Figure 61) may include the following sections, which you can expand and collapse by clicking and , respectively.

Figure 61. Item View Window: Rubric & Resources Tab



- **Details**: May provide the following information:
  - **Topic**: Skill area to which the item belongs.
  - **Difficulty**: Indicates whether the item is intended to be easy, moderate, or difficult.
  - **Content Alignment**: Describes the standard to which the item is aligned.
- **Resources**: Provides links to any exemplars or training guides available for the item.
- **Rubric**: Displays the criteria used to score the item. This section may also include a score breakdown, a human-readable rubric, or an exemplar, which provides an example of a response for each point value.
- **Frequency Distribution of Student Responses**: The table in this section provides a breakdown of how many students in the school earned each possible point value available for a fixed-form test item.

How to View Items with and Without the Students' Visual Settings

When viewing items with students' responses, you may or may not want to see the items exactly the way the students saw them on the test. For example, some students' tests are set to use large fonts, different color contrast, or Spanish.

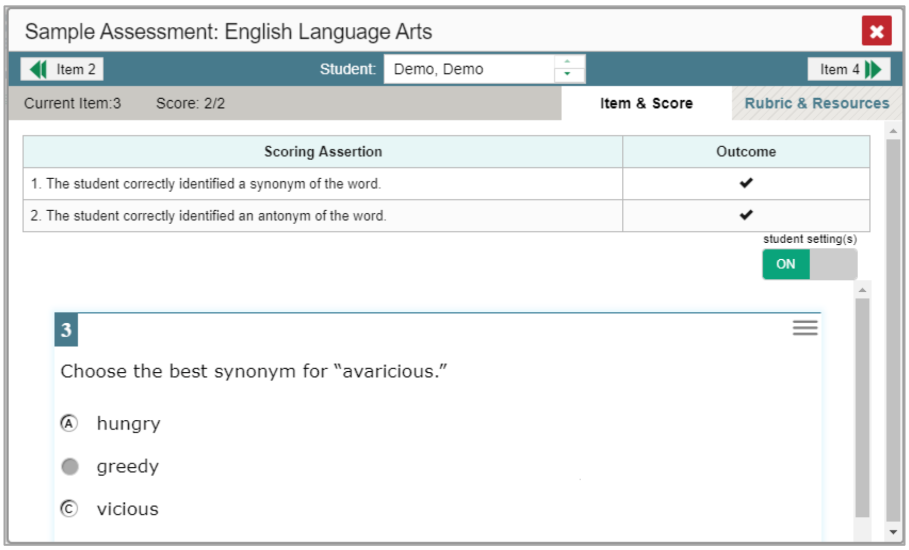
1. From the **Features & Tools** menu, select **Set Student Setting on Item View**. The ***Set Student Setting on Item View*** window appears.

2. Select **Yes** to show students' visual settings on all items or **No** to hide them.

3. Click **Save**.

You can also show or hide visual settings on a per-item basis. To do so, click the toggle at the upper right of the item you're viewing (see Figure 62). This action has no effect on your global setting.

Figure 62. Item View Window: Item & Score Tab with Student Response



What It Means When a Student Response Contains Highlighted Text

When a student's text response contains too much text copied from the item prompt and a condition code of Insufficient Original Text to Score has been applied, the copied portion is automatically highlighted.

How to Navigate to Other Items from the Item View Window

Use the buttons labeled with the previous and next item numbers at the upper corners of the ***Item View*** window.

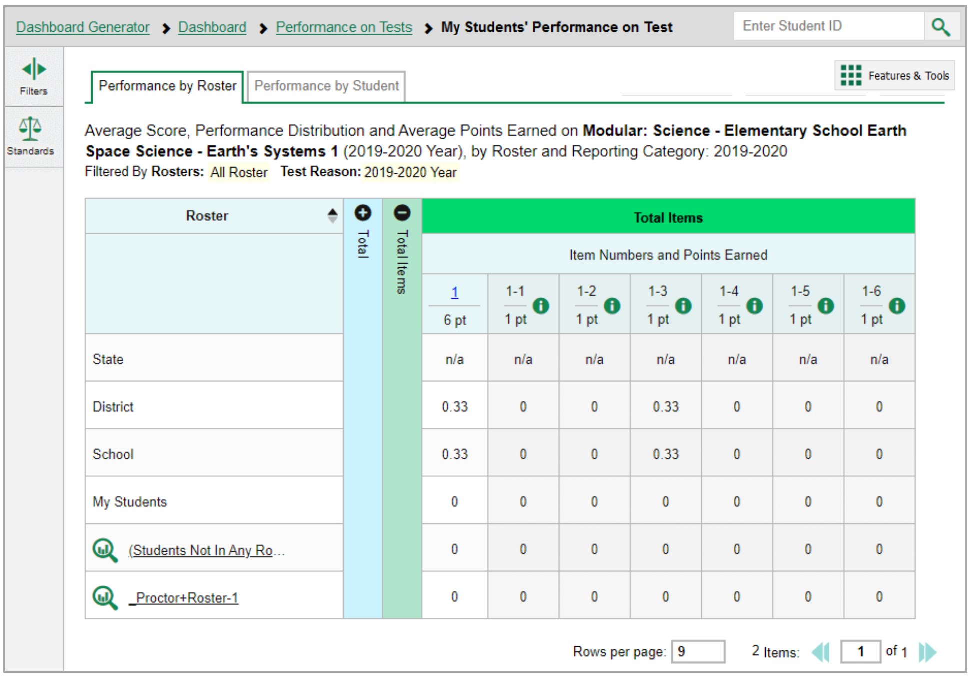
How to View Another Student's Response to the Current Item

If you have accessed the student's response from a report showing multiple students, you can click the arrows beside the *Student* field Demo Student at the top of the window. The students are listed in the same order in which they are sorted in the report.

What It Means When Items Are Labeled “1-1”, “1-2”, and So On (MSA Only)

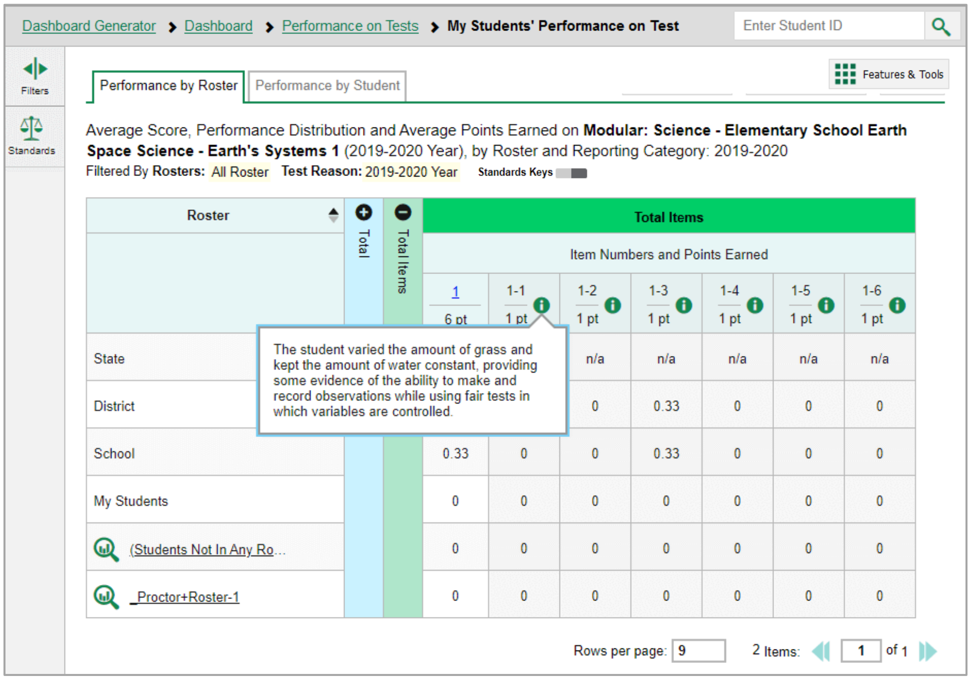
Those are sub-items belonging to an item cluster. Clusters are broken down into sub-items because they have multiple scoring assertions. Each sub-item has its own column to the right of the main item column. Sub-items are labeled “[item number]-[sub-item number]”, for example, “1-1”, “1-2”, “1-3”, as in Figure 63. This is applicable to the Montana Science tests.

Figure 63. My Students’ Performance on Test Report: Performance by Student Tab with Expanded Total Items Section



To view a scoring assertion, click the more information button to the right of the sub-item number, as in Figure 64.

Figure 64. My Students' Performance on Test Report: Performance by Student Tab with Expanded Total Items Section and Sub-Item Assertion



What It Means When an Item Score Reads "n/a"

You may sometimes see "n/a" instead of a score for an item. In some cases, the student did not respond to the item, or the item was not included in that form of the test.

How to Score Items on Interims (SBAC Only)

The Reporting System allows authorized users to score certain items on interim tests.

- Some items that require hand scoring arrive in the Reporting System without any scores. For example, all short answer items require hand scoring. If a test contains unscored items, its performance data are excluded from your reports until an authorized user scores all the unscored items in at least one opportunity of that test.

- Other items arrive in the Reporting System with automated scores suggested by the machine scoring system. Authorized users can override these scores if necessary. For example, all full write items have machine-suggested scores that can be overridden.

How to Score Unscored Items

For a student's test performance to be reported, you need to enter scores for any hand-scored item responses on that test.

When you have tests with unscored items, a **Tests To Score** notification appears in the banner.

1. In the banner, click **Tests To Score**.
2. If the ***Select School*** page appears, make a selection and click **Continue**.
3. On the scoring ***Dashboard*** (Figure 65), click the name of the test you wish to score. The ***Test Scoring*** page appears (Figure 66), displaying a list of students and items awaiting scoring for the selected test.

Figure 65. Scoring Dashboard

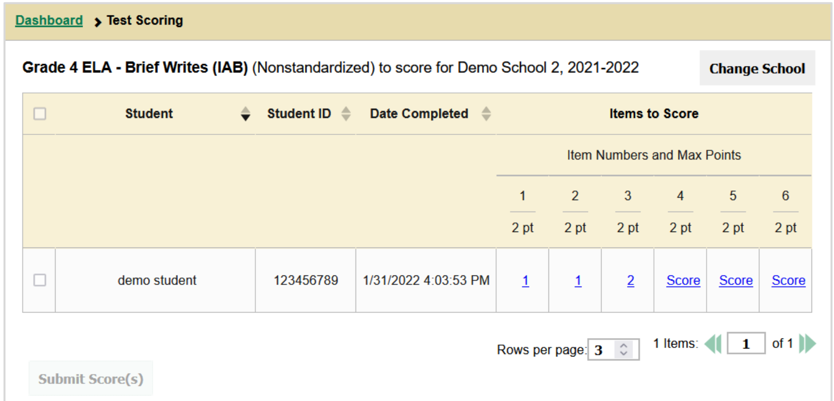
Dashboard

Change School

Assessments to score for Demo School 2, 2021-2022

Assessment Name	Test Reason	Student Count	Items to Score	Items to Submit	Last Date Completed
Grade 5 MATH - Performance Task - Turtle Habitat (IAB)	Nonstandardized	1	4	0	02/15/2022
Grade 5 ELA - Brief Writes (IAB)	Nonstandardized	1	4	2	02/15/2022
Grade 4 ELA - Performance Task (ICA)	Standardized/Benchmark	1	2	0	02/14/2022
Grade 4 ELA - Brief Writes (IAB)	Standardized/Benchmark	1	2	4	02/14/2022
Grade 3 ELA - Read Literary Texts (IAB)	Nonstandardized	1	1	0	01/31/2022

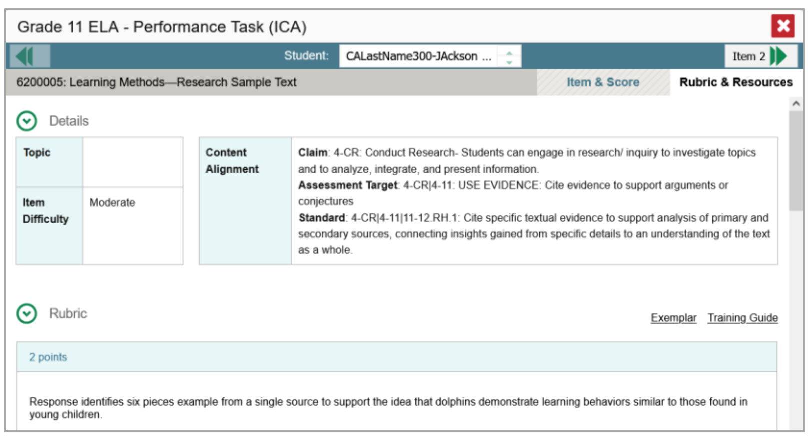
Figure 66. Test Scoring Page



4. To enter scores for an item response, click the **Score** link for the required item in the required student's row. The scoring window opens to the **Item & Score** tab.

5. *Optional*: You can open the **Rubric & Resources** tab (Figure 67) to review the item's scoring rubric and any other available resources, such as an exemplar and training guide for scoring the item.

Figure 67. Item View Window: Rubric & Resources Tab



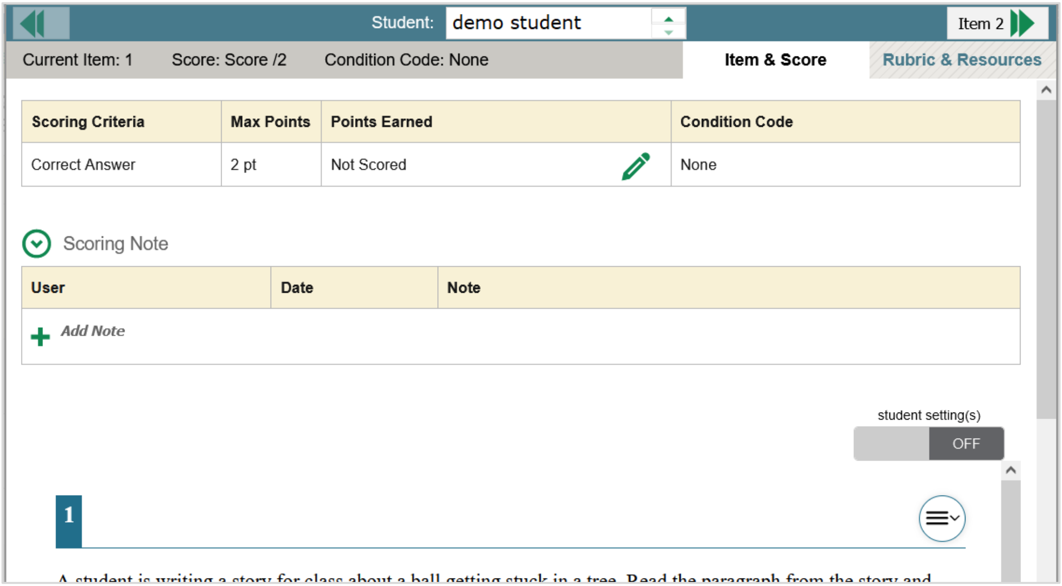
6. On the **Item & Score** tab (see Figure 68), click the edit button in the Scoring Criteria table at the top of the window.

7. Review the student's entered response and do one of the following:

   - If the student's response is scorable, select the appropriate score from the drop-down menu in the Points Earned column.

   - If the student's response cannot be scored for any reason, select the appropriate condition code from the drop-down in the Condition Code column. For more information about condition codes, see Condition Codes.

Figure 68. Item View Window: Item & Score Tab



8. If the item has multiple scoring criteria, repeat step 7 for each criterion.
   - When assigning condition codes to multiple scoring criteria, note that some related criteria may require the codes to be the same.
9. Click **Save**.
10. *Optional*: If you wish to provide an explanation for why you chose the given score or condition code, click **Add Note** in the *Scoring Note* section, then enter your comment and click **Save**.
11. To continue scoring items, do one of the following:
    - To view another unscored item for the same student, use the buttons labeled with the previous and next item numbers at the upper corners of the scoring window.
    - To view the same unscored item for another student, click the up or down arrows on the right side of the *Student* field Demo Student at the top of the window.
    - To return to the ***Scoring Mode*** window and select another item manually, close the ***Item View*** window using the close button at the upper-right corner.
12. Repeat steps 5–11 until you have entered scores for all the unscored items for the test.
13. *Optional*: If you need to change the entered score for an item response, you can select the score link for that item on the ***Test Scoring*** page and repeat steps 5–11.

Once you have entered scores for all the unscored items in a test opportunity, you may submit that opportunity for processing.

Note that for some items, two Scoring Criteria tables appear, with the top one allowing you to set scores and the bottom one displaying transformed scores (see Figure 69). When you set a score, the new scores are automatically transformed, and the transformed scores are what will appear in reports. You will need to refresh or navigate away from the item before you can view the transformed scores, and there may be a delay before they appear.

Figure 69. Item View Window: Item & Score Tab: Second Scoring Criteria Table with Transformed Scores

Transformed Points Earned and Condition Codes

Scoring Criteria	Max Points	Points Earned	Condition Code
Conventions	2	2	None
Organization, Purpose, Evidence and Elaboration	4	Condition Code Selected	Off Purpose
Overall	6	2	None

Full write items are scored using three dimensions: Conventions, Evidence/Elaboration and Organization/Purpose for a total of ten points. For test scoring purposes,

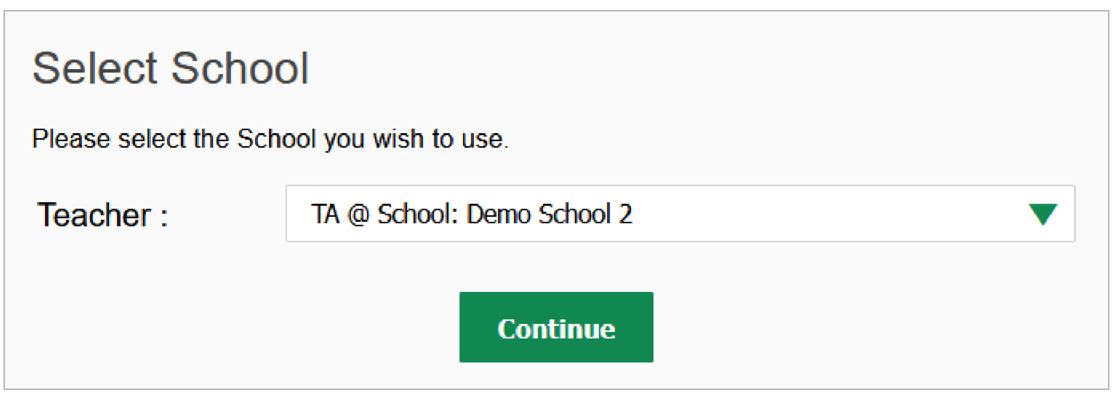
To return to reports, click the close button at the upper-right corner.

How to Change the Selected School

Scoring mode allows you to score the item responses for only one school's test opportunities at a time. If you are associated with multiple schools, you may change the selected school in order to score the test opportunities from other schools.

1. To change the selected school, click **Change School** in the top-right corner of the ***Dashboard***. The ***Select School*** page appears.

Figure 70. Select School Page



2. From the **Teacher** drop-down menu, select the school whose student responses you wish to score.

3. Click **Continue**. The ***Dashboard*** page appears, displaying the tests available for the selected school.

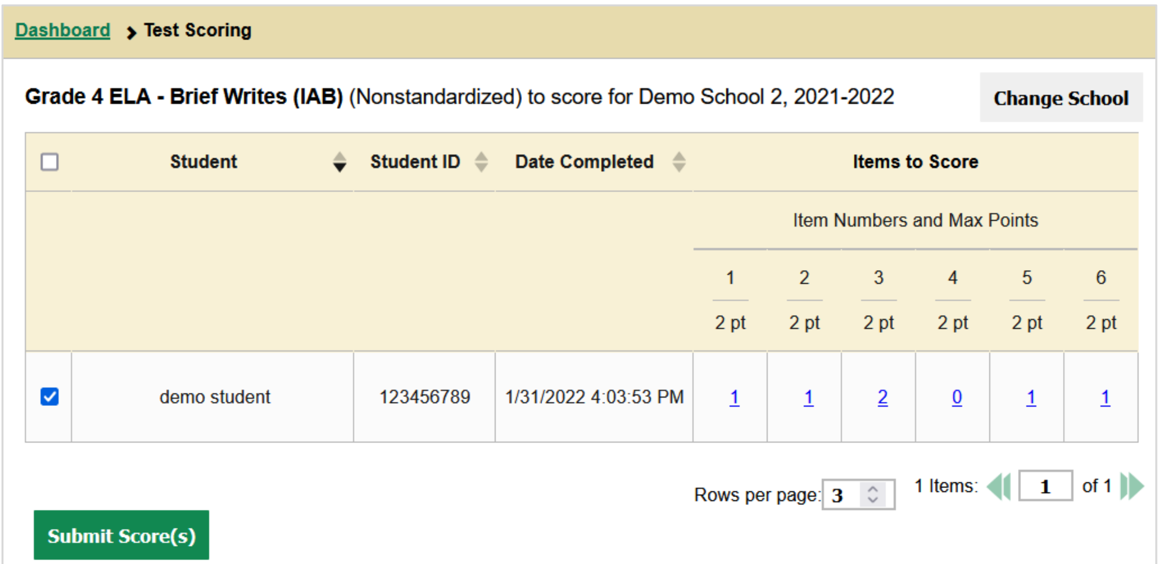
How to Submit Scored Test Opportunities for Processing

In order for a test opportunity's scores to be reported, you will need to submit that opportunity for processing. You may only submit an opportunity once you have entered scores or condition codes for every hand-scored item on the test.

1. To submit scored opportunities for processing, navigate to the ***Test Scoring*** page (see Figure 71) and mark the checkbox for each opportunity you wish to submit.

    - You can mark the checkbox in the top-left corner to select all fully scored opportunities at once.

Figure 71. Test Scoring Page: Submitting an Opportunity



2. Click **Submit Score(s)** in the bottom-left corner of the page.

3. In the confirmation window that pops up, click **Continue**. The selected opportunities will be submitted for processing and reporting and removed from scoring.

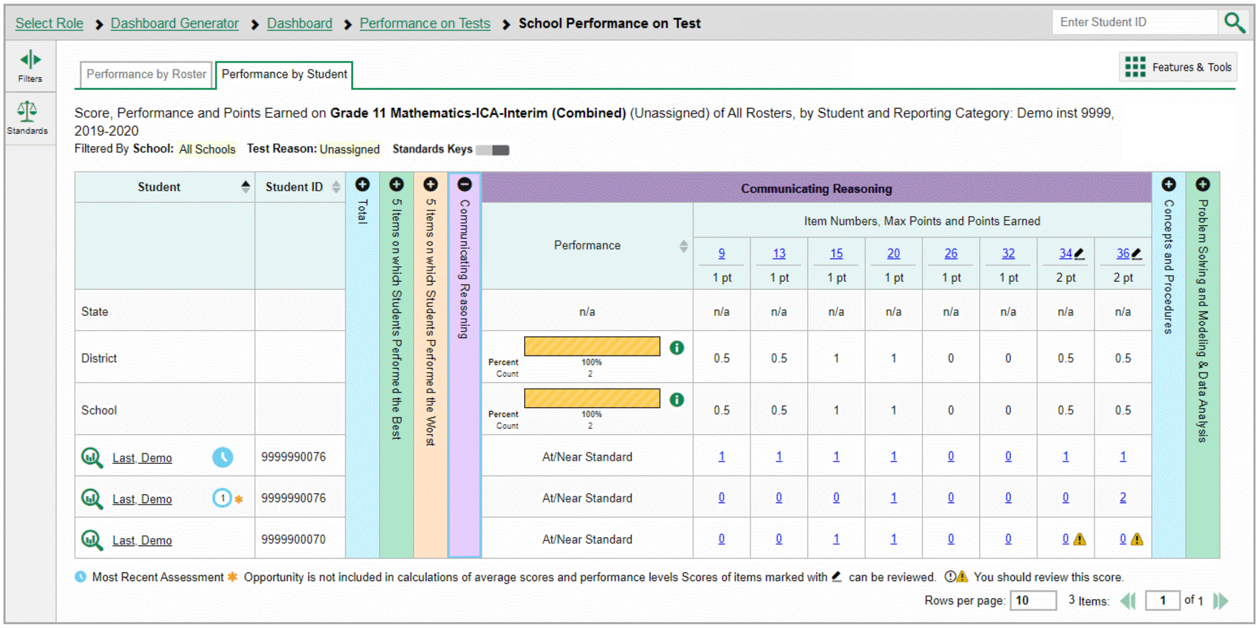
To return to reports, click the close button at the upper-right corner. You can still modify the item scores on that test directly from the reports by following the procedure in the next section (How to Modify Scores for Items).

How to Modify Scores for Items

You can modify scores for some items directly from the ***Item View*** window.

Reports display a pencil icon ✎ in the column header for each item with a modifiable score (see Figure 72). When a machine-suggested score has a low confidence level, or when a condition code of Insufficient Text or Non-Scorable Language has been assigned by machine, ⚠ displays next to the score. It is highly recommended that you review items flagged with this icon.

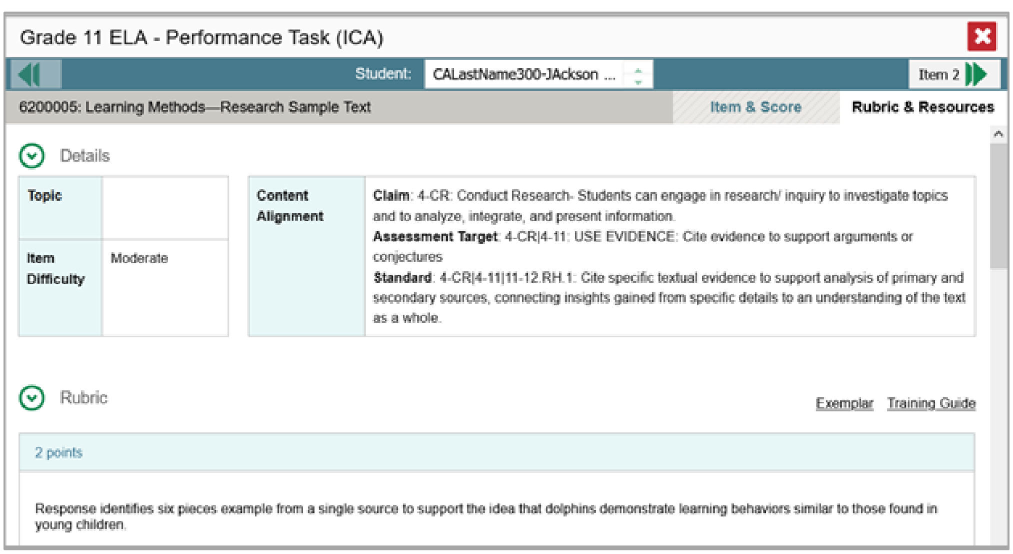
Figure 72. School Performance on Test Report: Performance by Student Tab with Expanded Reporting Category Section



1. On a report with modifiable scores, click the item score link in the student's row of the report. The ***Item View*** window opens.

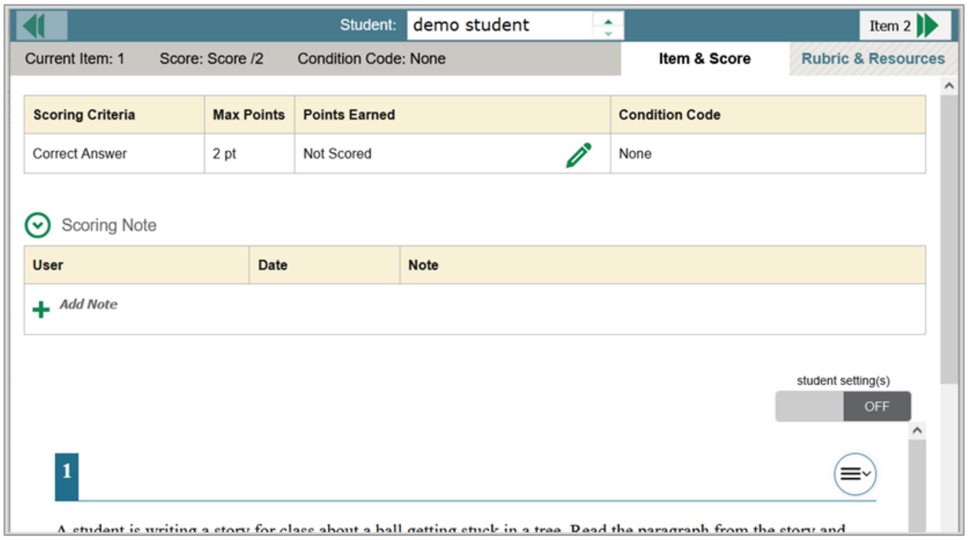
2. *Optional*: You can open the **Rubric & Resources** tab (Figure 73) to review the item's scoring rubric and any other available resources, such as an exemplar and training guide for scoring the item.

Figure 73. Item View Window: Rubric & Resources Tab



3. On the **Item & Score** tab (see Figure 74), review the student's entered response and click in the Scoring Criteria table at the top of the window. The Points Earned and Condition Code columns become editable.

Figure 74. Item View Window: Item & Score Tab



4. Do one of the following:

   - To enter a score for the response, select a numerical score from the **Points Earned** drop-down list.

   - To assign a condition code to the response, select one from the **Condition Code** drop-down list.

5. If the item has multiple scoring criteria, repeat step 4 for each criterion.

6. Click **Save**.

7. *Optional*: If you wish to provide an explanation for why you chose the given score or condition code, click **Add Note** in the *Scoring Note* section, then enter your comment and click **Save**.

8. To continue modifying scores, do one of the following:

   - To view another item for the same student, use the buttons labeled with the previous and next item numbers at the upper corners of the ***Item View*** window.

   - To view the same item for another student, use the up or down arrow buttons on the right side of the *Student* field Demo Student at the top of the ***Item View*** window.

The performance data in the test results update automatically when you close the ***Item View*** window.

Note that for some items, two Scoring Criteria tables appear, with the top one having modifiable scores and the bottom one displaying transformed scores, as in Figure 75. When you modify a score, the new scores are automatically transformed, and the transformed scores are what will appear in reports. You will need to refresh or navigate away from the item or the report before you can view the transformed scores, and there may be a delay before they appear.

Figure 75. Item View Window: Item & Score Tab: Second Scoring Criteria Table with Transformed Scores

Transformed Points Earned and Condition Codes

Scoring Criteria	Max Points	Points Earned	Condition Code
Conventions	2	2	None
Organization, Purpose, Evidence and Elaboration	4	Condition Code Selected	Off Purpose
Overall	6	2	None

Full write items are scored using three dimensions: Conventions, Evidence/Elaboration and Organization/Purpose for a total of ten points. For test scoring purposes,

How to Set Up Interim Reports to Suit Your Needs

There are three ways of setting up your interim reports that are different from summatives. You can assign test reasons to interim test opportunities, filter them by test reason, and filter them by standard.

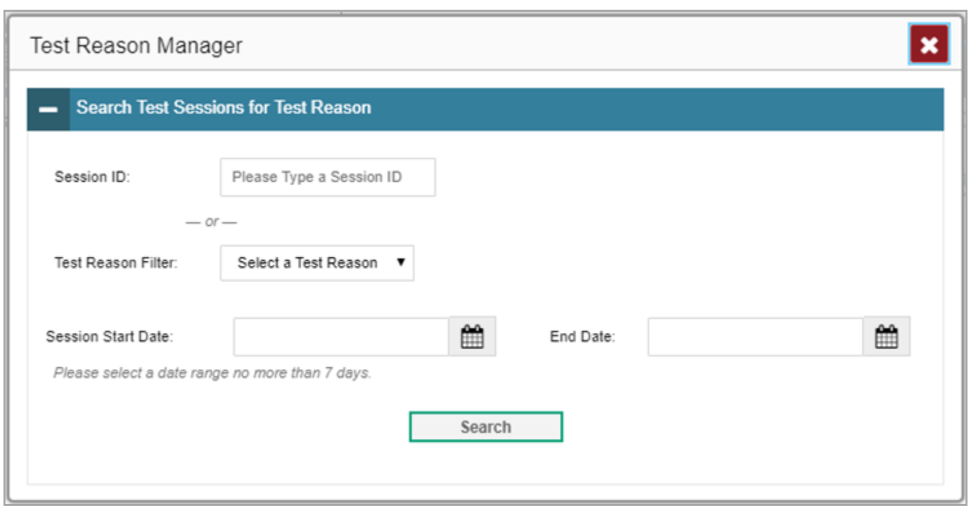
How to Assign Test Reasons (Categories) to Interim Test Opportunities

Test reasons are categories used to classify test opportunities for reporting purposes. They typically indicate the timeframe in which tests were taken, and they're a good way to organize tests into groups.

Test reasons should ideally be assigned in the Test Administration Site at the time of testing. However, you can use the Test Reason Manager in the Reporting System to assign a different test reason to an interim or benchmark test opportunity that was completed in the present school year. Summative test reasons cannot be reassigned.

1. From the **Features & Tools menu**, select **Manage Test Reasons**. The ***Test Reason Manager*** window opens (see Figure 76).

Figure 76. Test Reason Manager Window

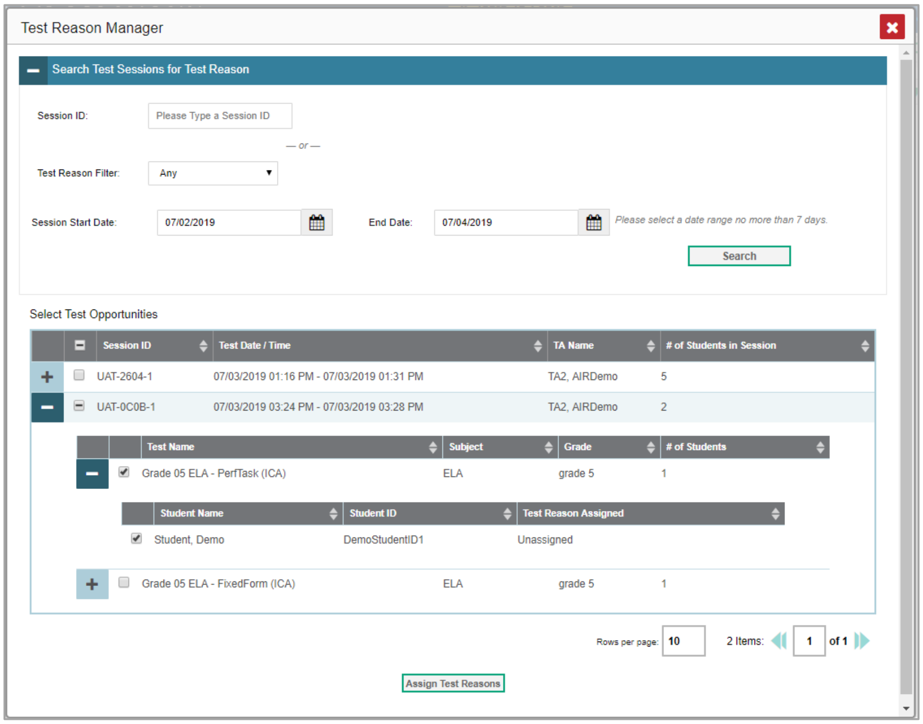


2. To search for the test opportunities you wish to categorize, do either of the following (see Figure 76):

- In the *Session ID* field, enter the session ID in which the opportunities were completed in TDS.
- Select the test reason associated with the opportunities you want to edit. Then select a range of dates during which the test session was administered. The date range cannot exceed seven days.

3. Click **Search**.
4. A list of retrieved test sessions appears in the section *Select Test Opportunities* (see Figure 77). You can click the  buttons to expand the list of tests in each session and the list of students who took each test (that is, individual test opportunities). To navigate through a long list, use the controls in the upper-right and lower-right corners.

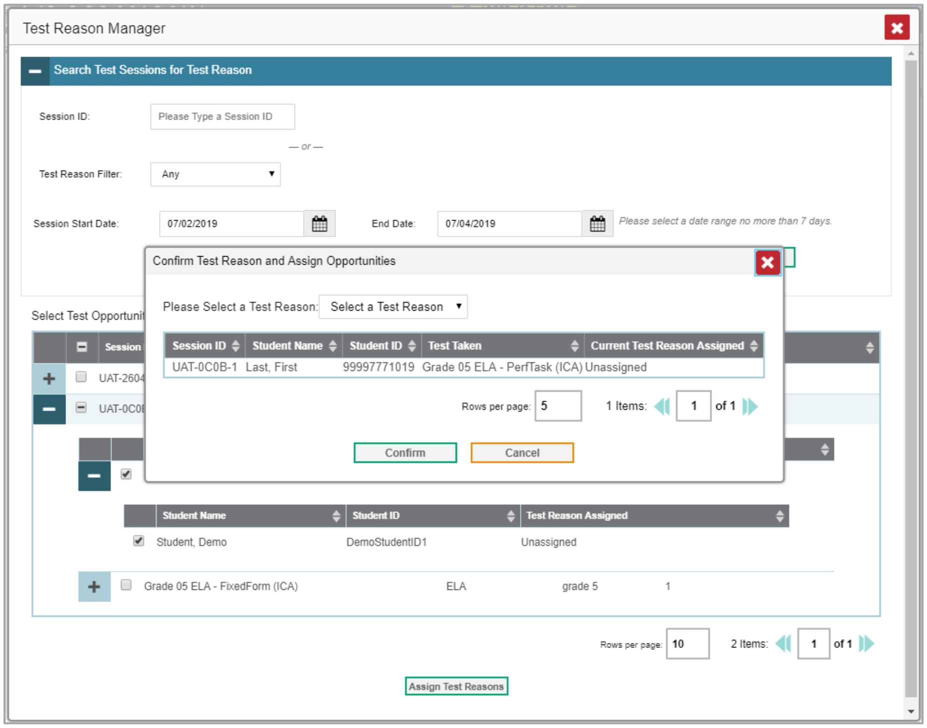
Figure 77. Test Reason Manager Window: Select Test Opportunities



5. Mark the checkboxes for each session, test, or opportunity that you wish to assign to a test reason.
6. Click **Assign Test Reasons** below the list of retrieved sessions.

7. In the window that appears (see Figure 78), select a new test reason to assign to the selected opportunities and click **Confirm**.

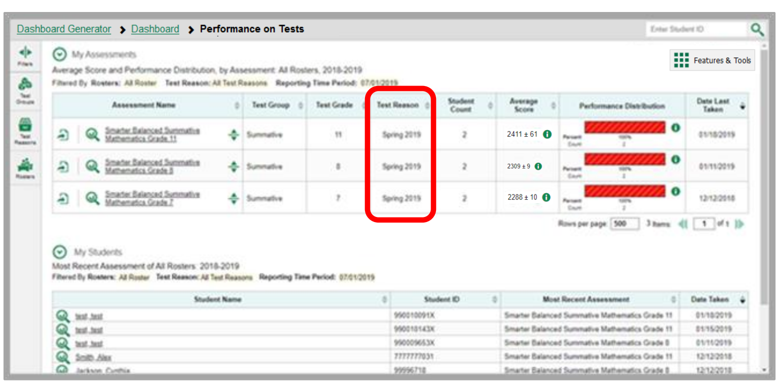
Figure 78. Confirm Test Reason and Assign Opportunities Window



How to Filter by Test Reason (Category)

Test reasons (shown in Figure 79) are categories used to classify test opportunities for reporting purposes. They typically indicate the timeframe in which interim and benchmark tests were taken, and they can be a good way to focus on specific groups of tests. For summative assessments, test reasons are simply test windows and are not useful.

Figure 79. Teacher View: Performance on Tests Report



When your test opportunities have test reasons, you can filter reports by a single test reason. For example, you may want to filter by Fall and look at ELA performance, then filter by Spring and see if students have improved on ELA material. If you don’t filter, you’ll see data for all different test reasons.

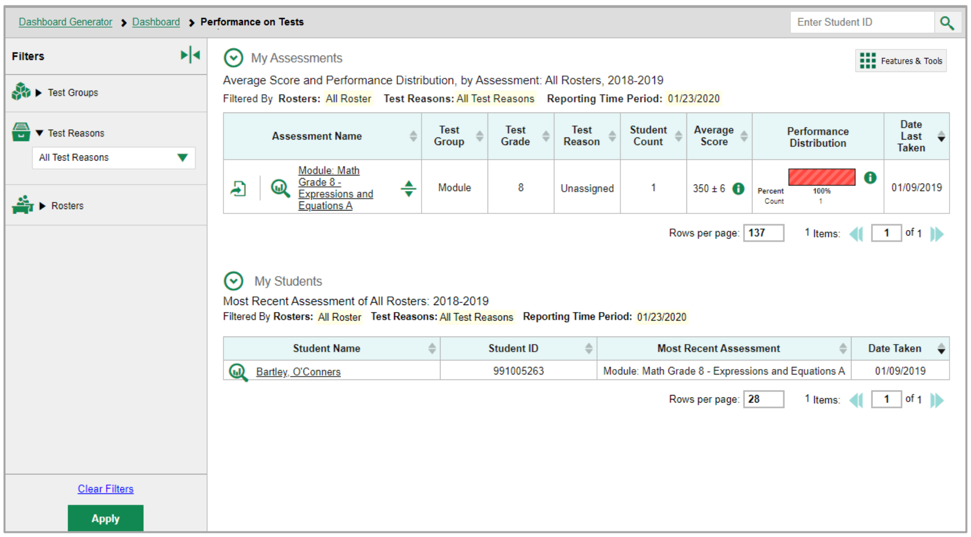
This will allow you to compare multiple test reasons side by side rather than a single test reason. You may find reports easier to understand when you're viewing only a single test reason.

The **Test Reasons** filter is available on the dashboards and Performance on Tests reports for teachers as well as for school- and district-level users.

1. On the left side of the dashboard or Performance on Tests report, click either the **Filters** panel expand button or the **Test Reasons** button. The **Filters** panel expands (see Figure 80).

2. Make a selection from the drop-down list in the **Test Reasons** section.

Figure 80. Teacher View: Performance on Tests Report with Expanded Filters Panel



3. Click **Apply**. The report updates to show only data for that test reason.

4. *Optional*: To revert all filters to their defaults, open the **Filters** panel again and click **Clear Filters**. Click **Apply**. Filters will also revert when you log out, switch user roles, or switch systems.

All the reports accessible from this page will be filtered the same way.

The row of filter details below the table header shows the test reason selected, if any.

How to Filter Item-Level Data on Interims by Standards and Clusters of Standards

An educational standard, sometimes called an assessment target, describes the skill the item measures. An example of a math standard is "At later grades, determine conditions under which an argument does and does not apply. (For example, area increases with perimeter for squares, but not for all plane figures.)"

You may want to see how your students performed on a particular standard or cluster of standards. In certain reports, you can filter by the standard to which items are aligned. That way you can view your students' performance in just one area of skill. Then you can switch filters to compare it with their performance in another skill. If you don't filter by standard, the reports will show results for all standards by default. You may find that switching between different sets of standard data and comparing them helps you understand students' abilities better.

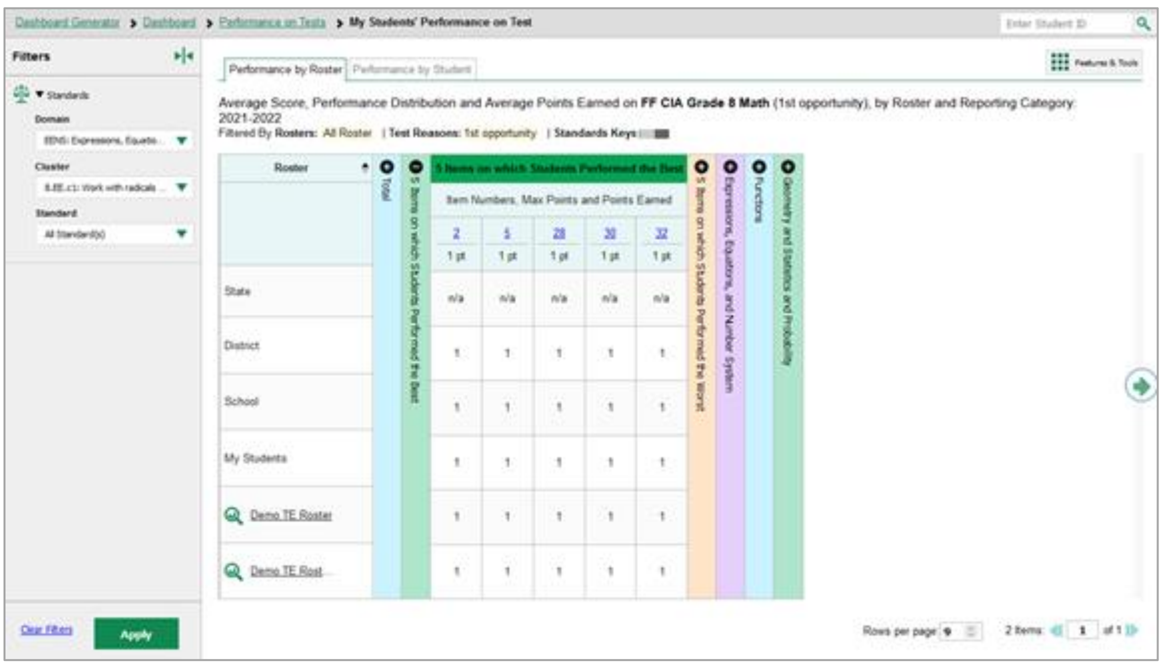
Standard filters are available in any report showing item-level data. The available standards vary by assessment.

1. On the left side of the page, click either the **Filters** panel expand button or the **Standards** button . The **Filters** panel expands (see Figure 81).

2. Use the drop-down list in the **Standards** section to select a cluster. An additional drop-down list appears.

3. *Optional*: Keep making selections from the drop-down lists as they appear.

Figure 81. My Students' Performance on Test Report: Performance by Roster Tab with Expanded Filters Panel



4. Click **Apply**. The affected report updates to show only the items that belong to the selected cluster or standard.

5. *Optional*: To revert all filters to their defaults, open the filters panel again and click **Clear Filters**. Click **Apply**. Filters will also revert when you log out, switch user roles, or switch systems.

All the reports accessible from this page will be filtered the same way.

The row of filter details below the table header specifies the standards selected, if any.

How to Export and Print Data on Interims

You can export or print any report you see in the Reporting System. Some reports on individual tests can be exported directly from the Performance on Tests report. You may want to export or print to save a snapshot of data to consult later, or to share data. Different options will be available depending on the report you are viewing. Some interim and benchmark reports can be exported with item-level data.

How to Export or Print a Report You're Viewing

1. Select the print button from the **Features & Tools** menu, or in a Longitudinal Report window, from the upper-right corner. If there are multiple report tables on the page, multiple print buttons will appear.

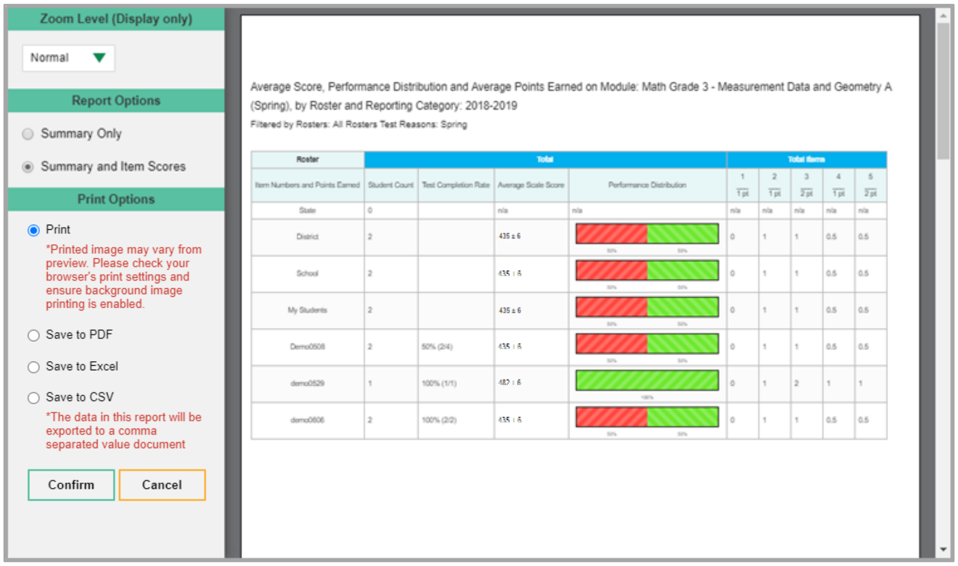
   A print preview page opens (see Figure 82).

   - To zoom in on the print preview, use the drop-down list under the *Zoom Level (Display only)* section. This setting affects the preview only.

Figure 82. Print Preview Page

2. If an aggregate report provides data for individual items, the *Report Options* section appears. Select either **Summary Only** or **Summary and Item Scores**. If you select the latter option, as in Figure 83, the printed report includes data for the individual assessment items. Printouts of the Student Performance on Test report always include item data if available.

Figure 83. Print Preview Page with Summary and Item Scores Option Selected



3. To print the report, select the **Print** radio button. To download it, select **Save to PDF**, **Save to Excel** (.xlsx), or **Save to CSV** (comma-separated values).

- *Optional*: If a printout or PDF is for a particular student, you can mark the **Include Items and Responses (takes extra time)** checkbox. The resulting report includes the actual items and the student's responses.
- If you selected **Save to PDF**, choose an option from the **Page Layout** drop-down list that appears.

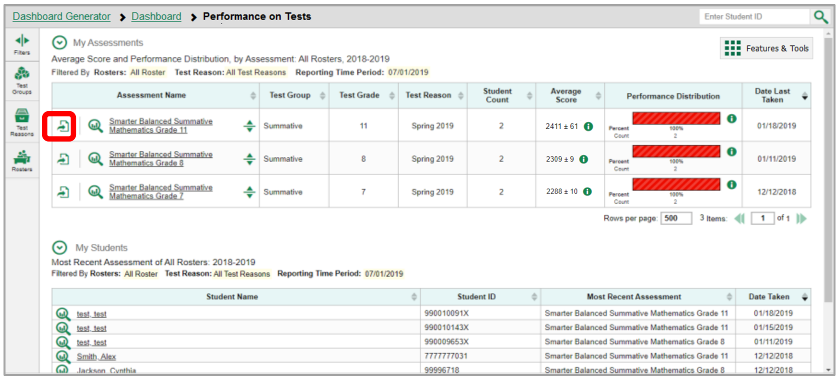
4. Click **Confirm**.

If you saved the report as a PDF, Excel, or CSV, the ***Secure Inbox*** window appears, displaying the generated report.

How to Export an Assessment Report Directly from the Performance on Tests Report

1. Click to the left of the name of the assessment whose report you wish to export (see Figure 84).

Figure 84. Teacher View: Performance on Tests Report



The ***Export Report*** window opens (see Figure 85 and Figure 86). The options in this window vary according to your user role.

2. Select which report to export for the assessment.

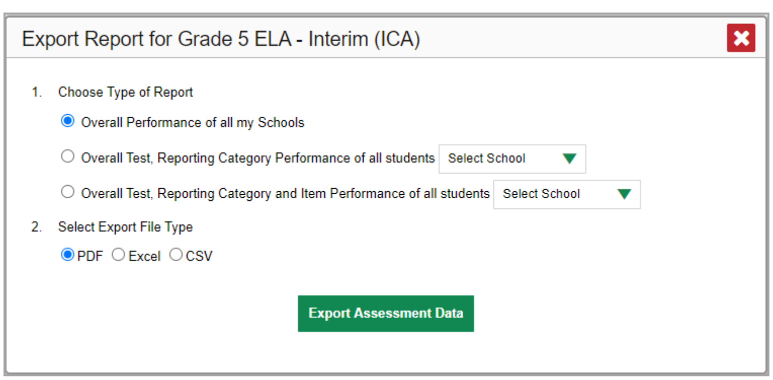
   - **District-level users:**

     – To export the district test results, mark the **Overall Performance of all my Schools** radio button.

     – To export school test results (excluding data for individual items), mark the **Overall Test & Reporting Category Performance of all my Students for [School Name]** radio button, then select a school from the drop-down list.

     – To export school test results (including data for individual items), mark the **Overall Test, Reporting Category and Item Performance of all my Students for [School Name]** radio button, then select a school from the drop-down list.

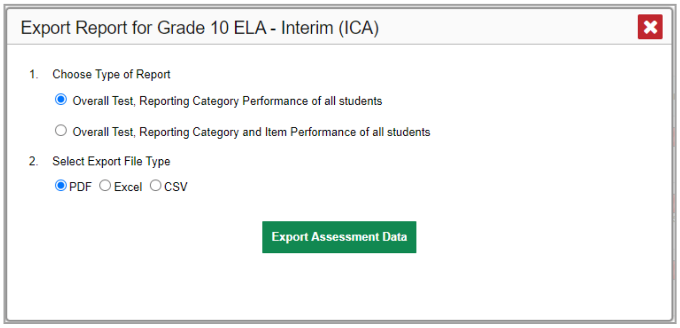
Figure 85. District-Level User View: Export Report Window



   - **School-level users and teachers:**

- To export results for all your associated students (excluding data for individual items), mark the **Overall Test, Reporting Category Performance of all students** radio button.
- To export results for all your associated students (including data for individual items), mark the **Overall Test, Reporting Category and Item Performance of all students** radio button.

Figure 86. Teacher View: Export Report Window



3. Choose from the **PDF**, **Excel**, and **CSV** formats.
4. Click **Export Assessment Data**. A confirmation window appears.
5. Click **Yes** to export or **No** to return to the ***Export Report*** window. When you've exported a file, the ***Secure Inbox*** window appears with the generated file available for download.

Appendix

Appendix sections are alphabetized for your convenience.

C

Class (Roster) Management

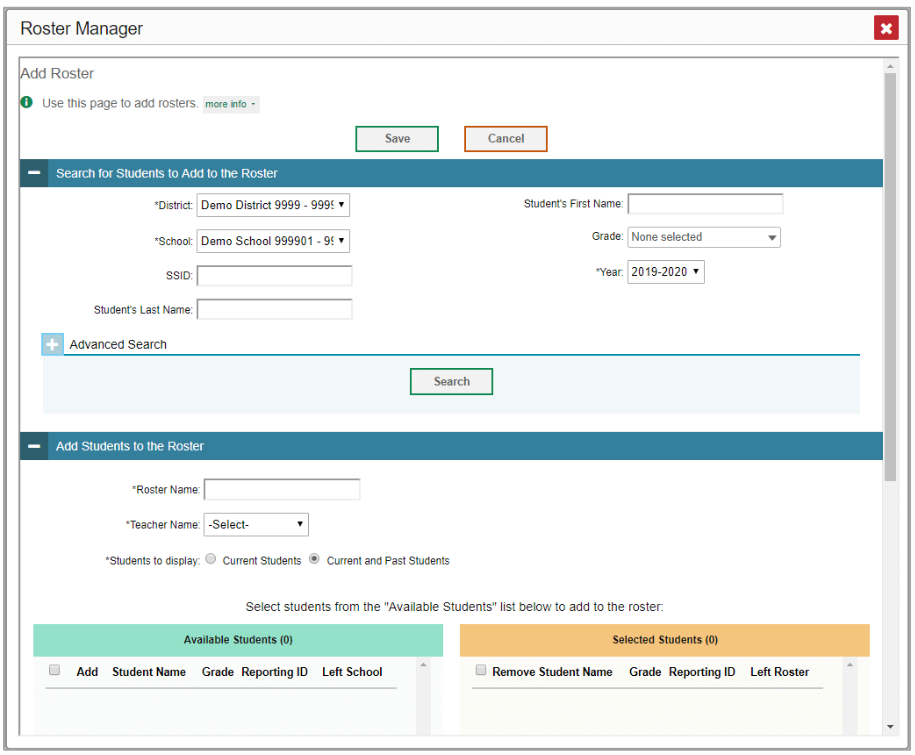
Teachers, school-level users, and district-level users can add, edit, and delete classes (rosters). Classes are a great way to organize students, allow teachers to view their students' performance, and allow other users to compare the performance of different classes.

How to Add a Class (Roster)

You can create new classes (rosters) from students associated with your school or district.

1. From the **Features & Tools menu**, select **Add Roster**. The ***Roster Manager*** window appears, showing the Add Roster form (see Figure 87).

Figure 87. Roster Manager Window: Add Roster Form



2. In the *Search for Students to Add to the Roster* panel (see Figure 87), do the following:

6. If you are a district-level user, then in the **School** drop-down list, select the school for the roster.

   b. *Optional*: In the *SSID*, *Student's First Name*, and/or *Student's Last Name* fields, enter information about a particular student you want to add.

7. *Optional*: In the **Enrolled Grade** drop-down list, select the grade levels for the students in the roster.

8. *Optional*: In the *Advanced Search* panel (see Figure 88), select additional criteria:

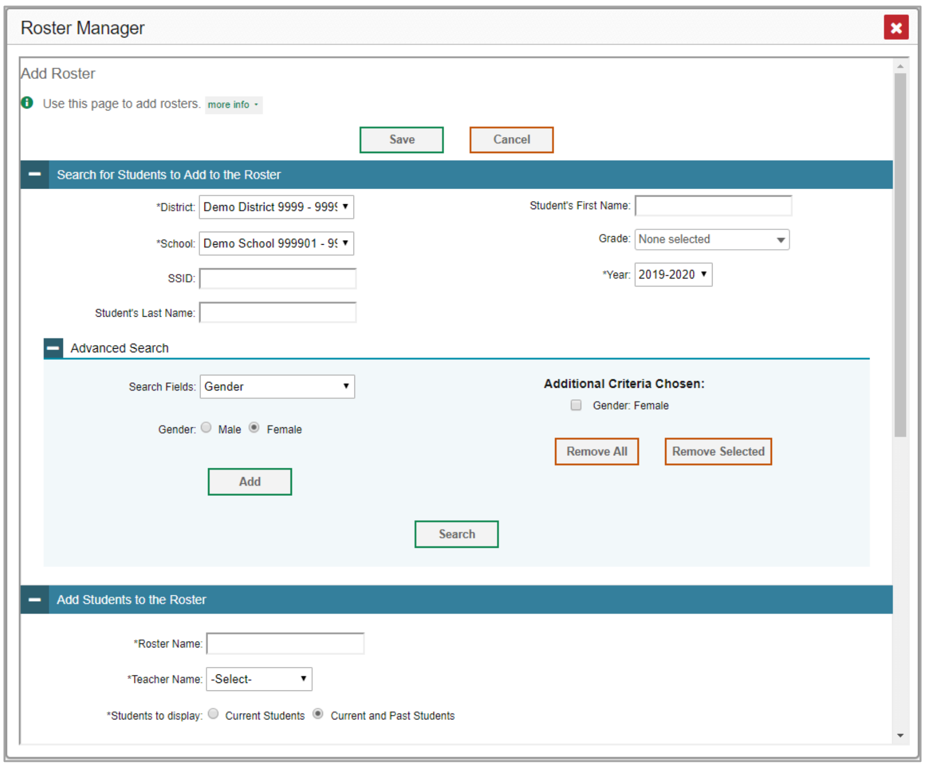
i. From the **Search Fields** drop-down list, select a criterion type. A set of related criteria for that criterion type appear.

ii. In the related fields, select the additional criteria.

iii. Click **Add**.

iv. *Optional*: To remove the added criteria, mark the checkboxes for those criteria and click **Remove Selected**. To remove all additional criteria, click **Remove All**.

Figure 88. Roster Manager Window: Add Roster Form with Advanced Search Panel in Use



9. Click **Search**. The *Add Students to the Roster* panel shows settings for the roster, a list of retrieved students (*Available Students*), and a blank *Selected Students* list.

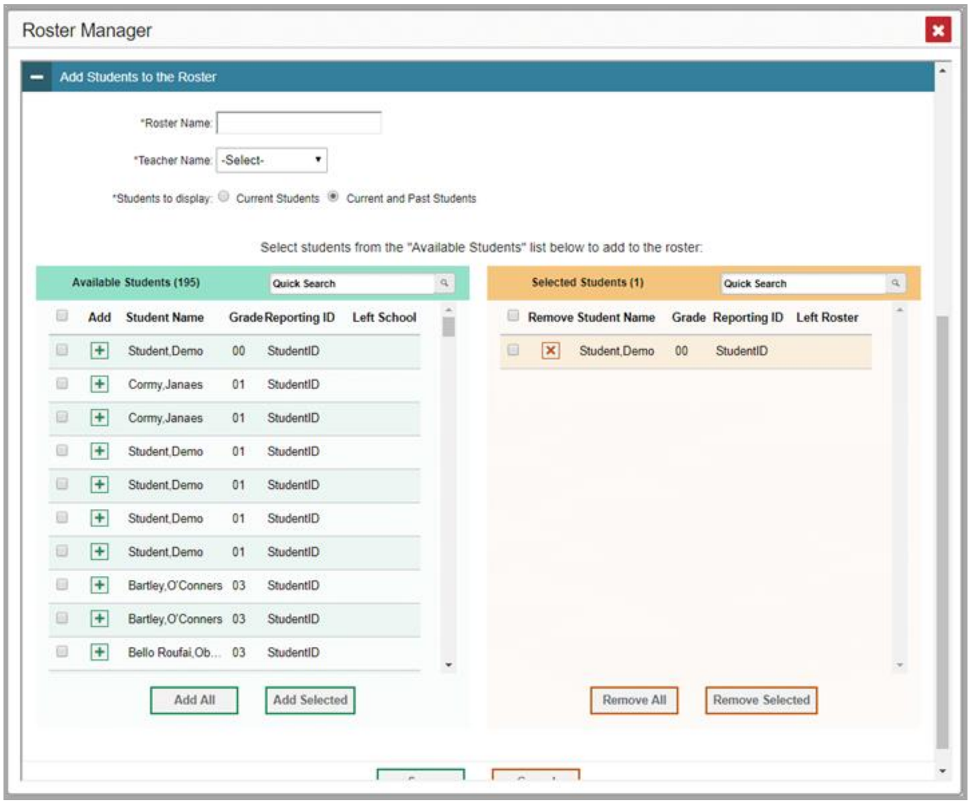
3. In the *Add Students to the Roster* panel (see Figure 89), do the following:

a. In the *Roster Name* field, enter the roster name.

10. From the **Teacher Name** drop-down list, select a teacher.

11. *Optional*: To include former students in the Add Roster form, mark the **Current and Past Students** radio button. The *Available Students* list will include students who have left the selected school.

Figure 89. Roster Manager: Add Roster Form Scrolled Down to Add Students to the Roster Panel



12. To add students, do one of the following in the list of available students:

- To move one student to the roster, click [+] beside that student's name.
- To move all the students in the *Available Students* list to the roster, click **Add All**.
- To move selected students to the roster, mark the checkboxes for the students you want to add, then click **Add Selected**.

13. To remove students, do one of the following in the list of students in this roster:

- To remove one student from the roster, click [×] beside that student's name.
- To remove all the students from the roster, click **Remove All**.
- To remove selected students from the roster, mark the checkboxes for the students you want to remove, then click **Remove Selected**.

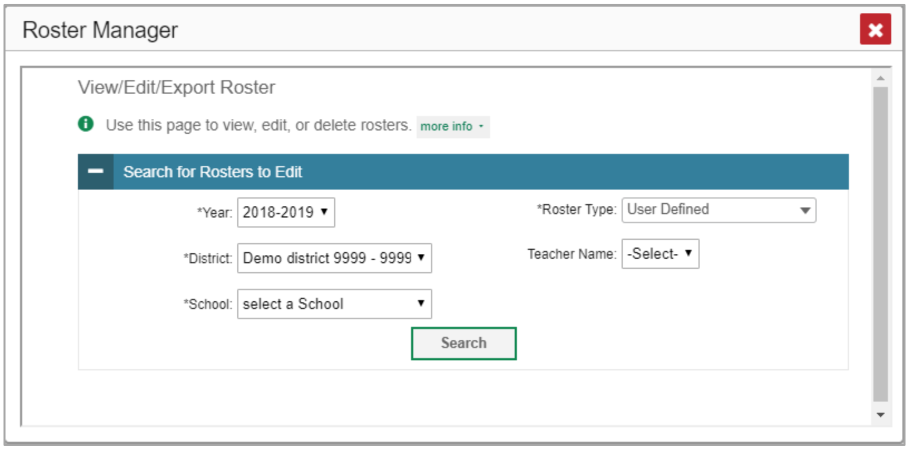
4. Click **Save**, and in the affirmation dialog box click **Continue**.

How to Modify a Class (Roster)

You can modify a class (roster) by changing its name, changing its associated teacher, adding students, or removing students.

1. From the **Features & Tools menu**, select **View/Edit Roster**. The ***Roster Manager*** window appears, showing the View/Edit/Export Roster form (see Figure 90).

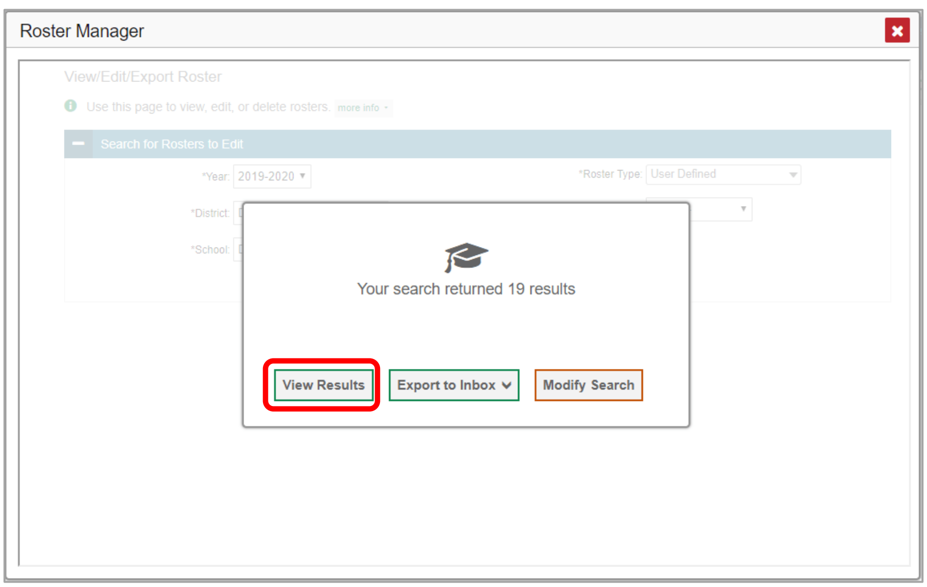
Figure 90. Roster Manager Window: View/Edit/Export Roster Form



2. In the *Search for Rosters to Edit* panel (see Figure 90), select the school year, school, and roster type for the roster you wish to edit. Optionally, select a teacher.

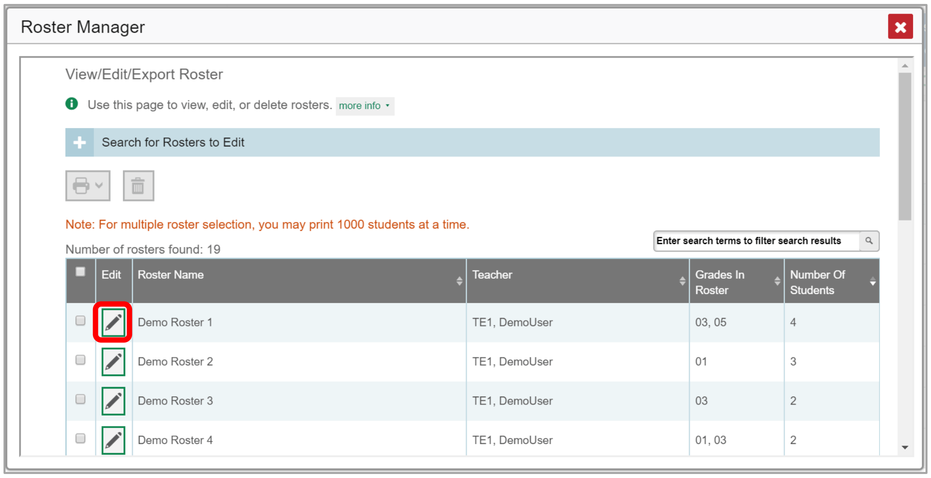
3. Click **Search**. A search results pop-up appears (see Figure 91). Click **View Results** to view the results in your browser.

Figure 91. Roster Manager Window: Search Results Pop-Up



4. A list of retrieved rosters is generated (see Figure 92).

Figure 92. Roster Manager Window: View/Edit/Export Roster Form Showing Retrieved Rosters



5. In the list of retrieved rosters, click  for the roster whose details you want to view. The ***View/Edit/Export Roster*** window opens.

6. *Optional*: To find students to add to the roster, use the *Search for Students to Add to the Roster* panel as follows:

a. If you are a district-level user, then in the **School** drop-down list, select the school for the roster.

b. *Optional*: In the SSID, *Student's First Name*, and/or *Student's Last Name* fields, enter information about a particular student you want to add.

14. *Optional*: In the Enrolled **Grade** drop-down list, select the grade levels for the students in the roster.

15. *Optional*: In the *Advanced Search* panel, select additional criteria:

i. From the **Search Fields** drop-down list, select a criterion type. A set of related criteria for that criterion type appear.

ii. In the related fields, select the additional criteria.

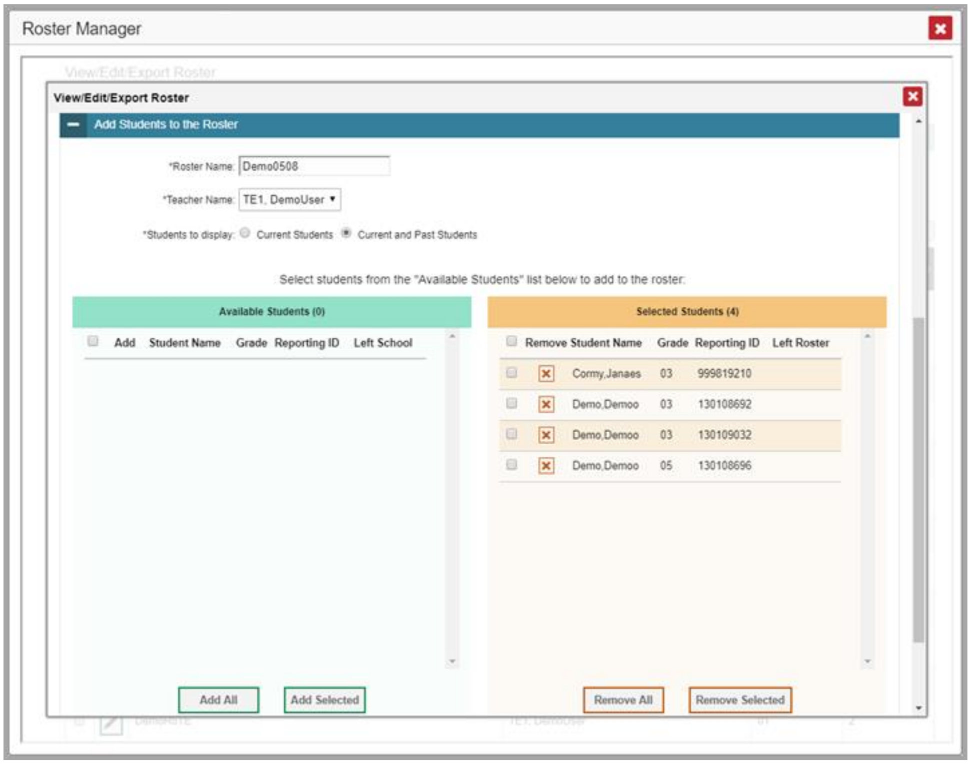
iii. Click **Add**.

iv. *Optional*: To remove the added criteria, mark the checkboxes for those criteria and click **Remove Selected**. To remove all additional criteria, click **Remove All**.

16. Click **Search**. The *Add Students to the Roster* panel shows settings for the roster, a list of retrieved students (Available Students), and a blank Selected Students list.

7. Scroll down to view the *Add Students to the Roster* panel, as in Figure 93.

Figure 93. Roster Manager Window: View/Edit/Export Roster Form Scrolled Down to the Add Students to the Roster Panel



8. *Optional*: In the *Add Students to the Roster* panel, do the following:

a. In the *Roster Name* field, enter a new name for the roster.

17. From the **Teacher Name** drop-down list, select the roster's new teacher.

18. *Optional*: To include former students in the Edit Roster form, mark the **Current and Past Students** radio button. The *Available Students* list will include students who have left the selected school, while the *Selected Students* list will include students who have left the roster.

19. To add students, do one of the following in the list of available students:

- To move one student to the roster, click [+] beside that student's name.
- To move all the students in the *Available Students* list to the roster, click **Add All**.
- To move selected students to the roster, mark the checkboxes for the students you want to add, then click **Add Selected**.

20. To remove students, do one of the following in the list of students in this roster:

- To remove one student from the roster, click beside that student's name.

- To remove all the students from the roster, click **Remove All**.
- To remove selected students from the roster, mark the checkboxes for the students you want to remove, then click **Remove Selected**.

9. At the bottom of the page, click **Save**, and in the affirmation dialog box click **Continue**.

How to Upload Classes (Rosters)

If you have many classes (rosters) to create, it may be easier to perform those transactions through file uploads. This task requires familiarity with composing comma-separated value (CSV) files or working with Microsoft Excel.

1. From the **Features & Tools menu**, select **Upload Rosters**. The ***Roster Manager*** window appears, showing the Upload Rosters: Upload page (see Figure 94).

Figure 94. Roster Manager Window: Upload Rosters: Upload Page

2. On the Upload Rosters: Upload page (see Figure 94), click **Download Templates** in the upper-right corner and select the appropriate file type (either **Excel** or **CSV**).

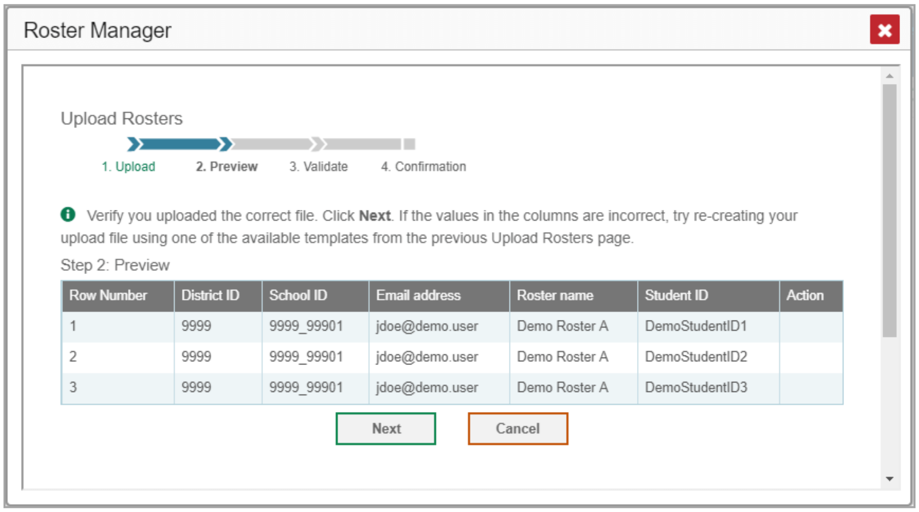
3. Open the template file in a spreadsheet application.

4. Fill out the template and save it.

5. On the Upload Rosters: Upload page, click **Browse** and select the file you created in the previous step.

6. Click **Next**. The Upload Rosters: Preview page appears (see Figure 95). Use the file preview on this page to verify you uploaded the correct file.

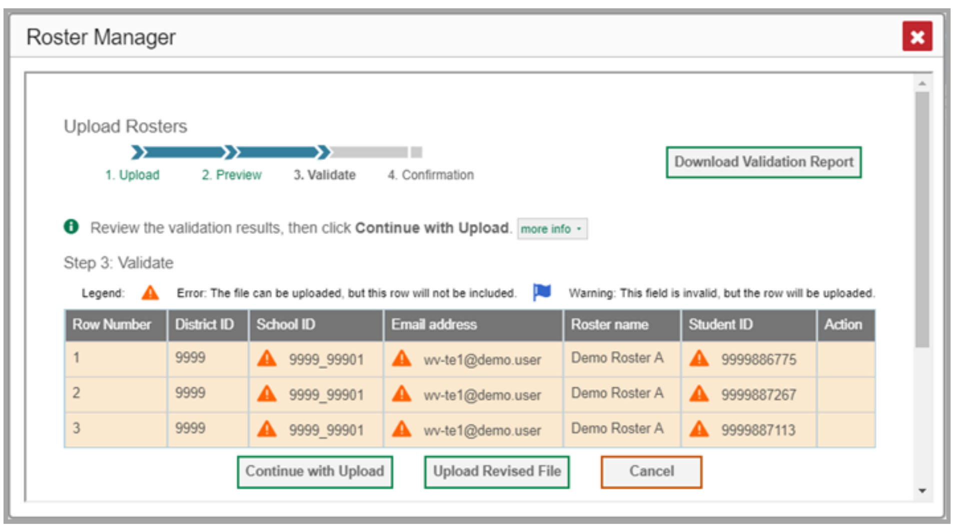
Figure 95. Roster Manager Window: Upload Rosters: Preview Page



7. Click **Next** to validate the file.

Any errors or warnings are displayed on the Upload Rosters: Validate page (see Figure 96). If a record contains an error, that record will not be included in the upload. If a record contains a warning, that record will be uploaded, but the field with the warning will be invalid.

Figure 96. Roster Manager Window: Upload Rosters: Validate Page



- *Optional*: Click the error and warning icons in the validation results to view the reason a field is invalid.
- *Optional*: Click **Download Validation Report** in the upper-right corner to view a text file listing the validation results for the upload file.

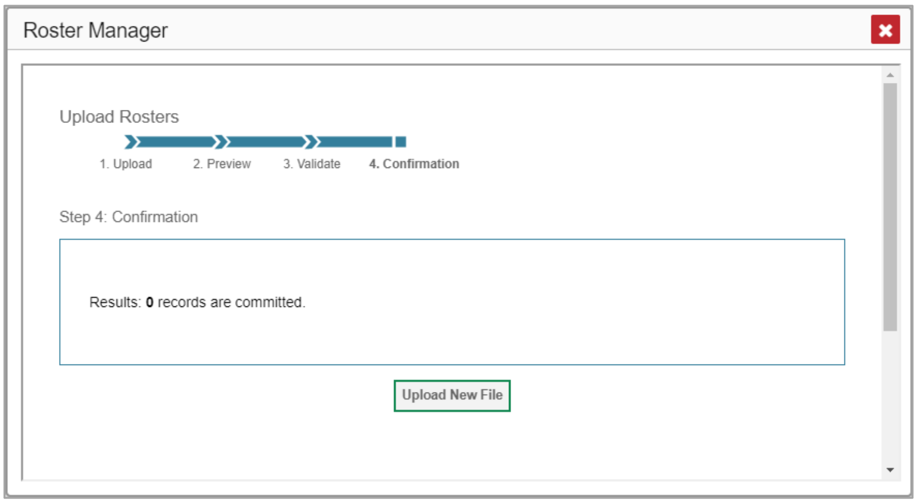
If your file contains a large number of records, the Reporting System processes it offline and sends you a confirmation email when it's complete. While the Reporting System is validating the file, do not press **Cancel**, as some records may have already started processing.

8. Do one of the following:

- Click **Continue with Upload** at the bottom of the page. The Reporting System commits those records that do not have errors. If there are too many errors, you won't be able to do this.
- Click **Upload Revised File** at the bottom of the page to upload a different file. Follow the prompts on the Upload Revised File page to submit, validate, and commit the file.

The Confirmation page appears (see Figure 97), displaying a message about how many records (rows) were committed.

Figure 97. Upload Rosters: Confirmation Page



9. *Optional*: To upload another roster file, click **Upload New File**.

Table 2 provides the guidelines for filling out the Roster template that you can download from the Upload Roster page.

Table 2. Columns in the Rosters Upload File

Element	Description	Valid Values
District ID*	District associated with the roster.	District ID that exists in TIDE. Up to 20 characters.
School ID*	School associated with the roster.	School ID that exists in TIDE. Up to 20 characters. Must be associated with the district ID. Can be blank when adding district-level rosters.
Email Address*	Email address of the teacher associated with the roster.	Email address of a teacher existing in TIDE.
Roster Name*	Name of the roster.	Up to 20 characters.
SSID*	Student's unique identifier within the district.	Up to 30 alphanumeric characters.

Element	Description	Valid Values
ACTION	Action to be taken on the student, either adding them to or deleting them from the roster. If blank, the student will be added.	`Add` or `Delete`.

*Required field.

Condition Codes

Table 3 provides an overview of the various condition codes that may be entered for a machine- or hand-scored item when a traditional score cannot be entered for the student's response.

Table 3. Condition Codes

Source of Code	Condition Code	Description
Human	Blank	• The student did not enter a response.
Human	Insufficient Text	• The student has not provided a meaningful response. Some examples: ▪ Random keystrokes ▪ Undecipherable text ▪ "I hate this test" ▪ "I don't know", "IDK" ▪ "I don't care" ▪ "I like pizza!" (in response to a reading passage about helicopters) ▪ Response consisting entirely of profanity • For ELA ICA Full Writes, use the "Insufficient Text" code for responses described above and also if ▪ The student's original work is insufficient to make a determination whether the student is able to organize, cite evidence/elaborate, and use conventions as defined in the rubrics. ▪ The response is too brief to make a determination regarding whether it is on purpose or on topic.
Human	Non-Scorable Language	• ELA/literacy: Language other than English. • Mathematics: Language other than English or Spanish.

Source of Code	Condition Code	Description
Human	Off Purpose	• For ELA ICA Full Writes only: ▪ A writing sample will be judged off purpose when the student has clearly not written to the purpose designated in the task. ▪ An off-purpose response addresses the topic of the task but not the purpose of the task. ▪ Note that students may use narrative techniques in an explanatory essay or use argumentative/persuasive techniques to explain, for example, and still be on purpose. ▪ Off-purpose responses are generally developed responses (essays, poems, etc.) clearly not written to the designated purpose.
Human	Off Topic	• For ELA ICA Full Writes only: ▪ A writing sample will be judged off topic when the response is unrelated to the task or the sources or shows no evidence that the student has read the task or the sources (especially for informational/explanatory and opinion/argumentative). ▪ Off-topic responses are generally substantial responses.
Machine	Blank	• The student did not enter a response.
Machine	Insufficient Text (Duplicated Text)	• The response contains a significant amount of text repeated over and over.
Machine	Insufficient Text (Too Few Words)	• The response contains too few words to be considered a valid attempt.
Machine	Insufficient Text (Copied Text from the Prompt)	• The response is largely composed of text copied from the prompt.
Machine	Insufficient Text (Refused to Answer)	• The response is a refusal to respond, in a form such as “idk” or “I don’t know.”
Machine	Non-Specific	• This condition code is assigned to machine-scored responses when TDS identifies that the response requires a condition code but cannot determine which specific condition code it requires.
Machine	Non-Scorable Language (Spanish Response)	• The response is in Spanish.

Source of Code	Condition Code	Description
Machine	Non-Scorable Language (Uninterpretable Language)	• The response is in a language other than English or Spanish.

H

Help

The Reporting System includes an online user guide.

How to Access the Online User Guide

In the banner, click **Help**. The guide opens in a pop-up window, showing the help page specific to the page you're on. For example, if you click **Help** while on the dashboard, you'll see the Overview of the Dashboard page.

I

Inbox

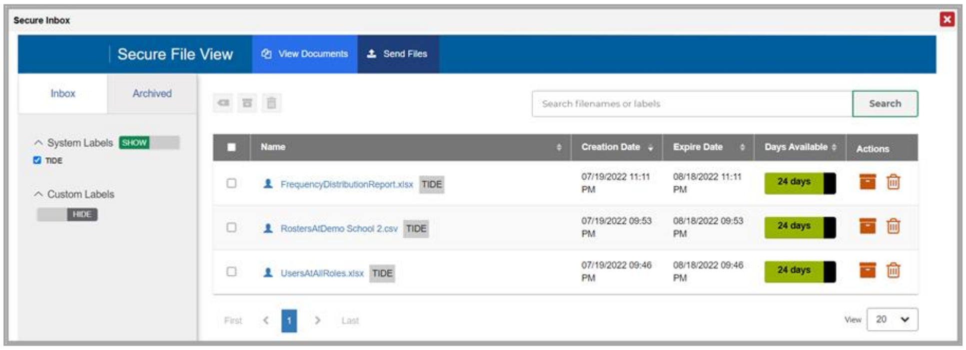
The Reporting System allows you to access a Secure Inbox feature that is integrated with other online assessment systems, such as TIDE, and accessible from your portal. The Inbox serves as a central repository for secure documents uploaded by administrators (such as state personnel) or shared between users, files exported by users, and hotline alerts.

Each user's Secure Inbox is personal to them and not shared among other users. Users can easily manage the files in their Inbox. The files are categorized into different tabs to allow users to view non-archived and archived files. Users can also search for files by keyword. Files are listed in the order in which they were created. The file creation and file expiration dates appear, if applicable, and the number of days remaining until a file expires is also displayed. By default, files are available for 30 days after being created. Users can archive or delete files as needed. Users can also share files by sending them to other users' Inboxes.

How to Access and Manage Files in the Inbox

1. In the banner, click **Inbox**. The ***Secure Inbox*** window appears (see Figure 98). By default, the Inbox window displays the **View Documents** tab.

Figure 98. Secure Inbox Window: View Documents Tab: Inbox Sub-Tab



2. Choose either of the available tabs (see Figure 98):

   - **Inbox**: Displays all files except those that have been archived. Includes columns for Creation Date, Expire Date, and Days Available.

   - **Archived**: Displays files that have been archived. Includes the same columns as the main **Inbox** tab.

3. *Optional*: To filter the files displayed, enter a search term in the text box in the upper-right corner and click . The search applies to both filenames and labels.

4. *Optional*: To hide or display system labels, click the System Labels toggle.

5. *Optional*: To hide files with a particular system label, clear the checkbox for that label.

6. *Optional*: To hide or display custom labels, click the Custom Labels toggle.

7. *Optional*: To hide files with a particular custom label, clear the checkbox for that label.

8. *Optional*: Do one of the following:

   - To download a file, click the name of the file.

   - To apply a custom label, follow these instructions:

     – To create a new custom label, mark the checkbox for any file, click the label button , enter a new custom label in the text box, and click **Save New Label**. Then apply it as described below.

     – To apply a custom label to a file, mark the checkbox for that file, click the label button , mark the checkbox for that label, and click **Apply Label**.

   - To archive a file, click .

   - To unarchive a file, click . The file is moved back to the main Inbox.

- To delete a file, click .

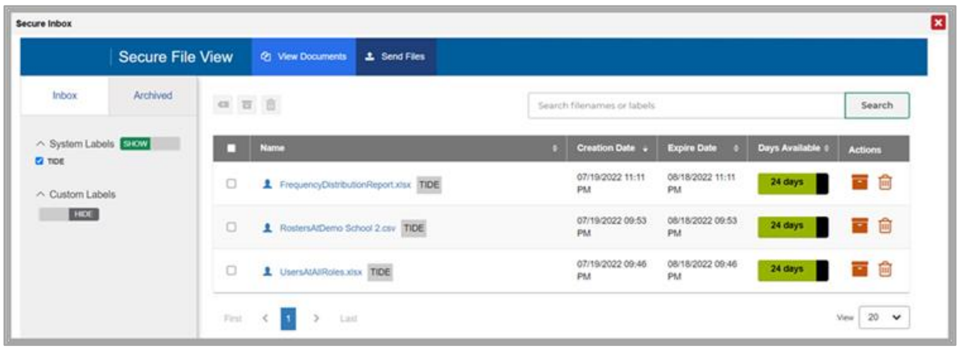
You cannot delete or archive secure documents uploaded to the Inbox by admin users.

How to Use the Inbox to Send Files to Other Users' Inboxes

You can send a file or files to individual recipients by email address or to groups of recipients by user role.

1. From the banner, select **Inbox**. The ***Secure Inbox*** page appears (see Figure 99). By default, the ***View Documents*** tab displays.

Figure 99. Secure Inbox Window: View Documents Tab: Inbox Sub-Tab



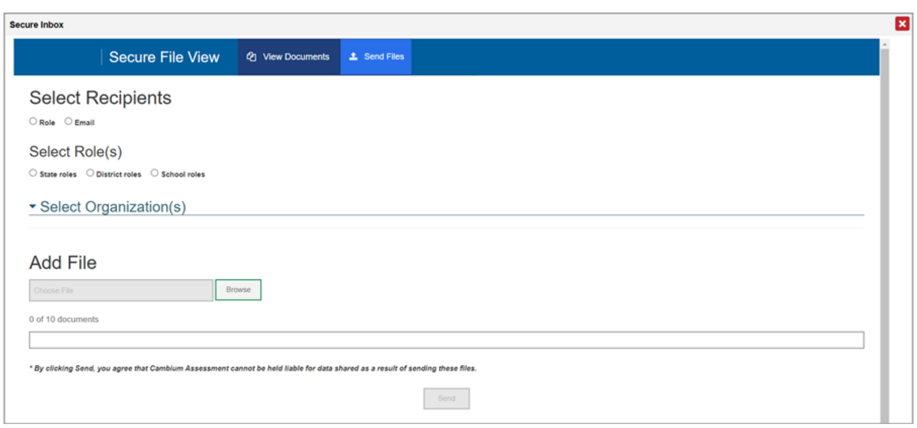
2. Select the **Send Files** tab. The ***Send Files*** page appears (see Figure 100).

3. In the *Select Recipients* field, do one of the following:

   - Select **Role** to send a file or files to a group of users by user role.
   - Select **Email** to send a file or files to a single recipient by email address.

   If you select **Email**, skip to step 7.

Figure 100. Secure Inbox Window: Send Files Tab



4. In the *Select Role(s)* field, select the role group to which you want to send a file or files. A drop-down list appears.

5. From the drop-down list, select the role(s) to which you want to send a file or files. You can choose **Select all** to send a file or files to all roles in the selected role group.

6. From the *Select Organization(s)* drop-down lists, select organizations that will receive the file(s) you send. These drop-down lists adhere to the user role hierarchy. For example, district-level users will be able to filter at their role level and below.

7. If you selected **Role** in step 3, skip this step. If you selected **Email**, enter the email address of the recipient to whom you wish to send a file or files.

8. To select a file or files to send, in the *Add File* field, select **Browse**. A file browser appears.

9. Select the file(s) you wish to send. You may send up to 10 files at once.

10. Select **Send**.

L

Login Process

This section describes how to log in to the Reporting System.

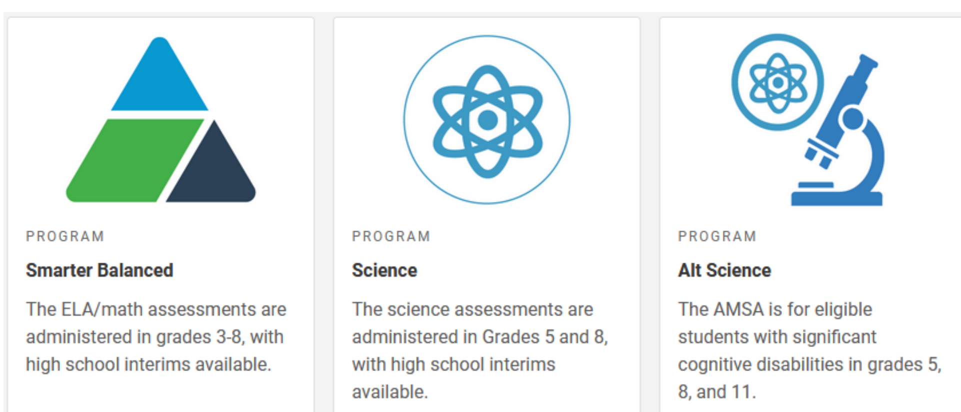
Do not share your login information with anyone. All Assessment Program systems provide access to student information, which must be protected in accordance with federal privacy laws.

How to Log In to the Reporting System

1. Navigate to the portal.

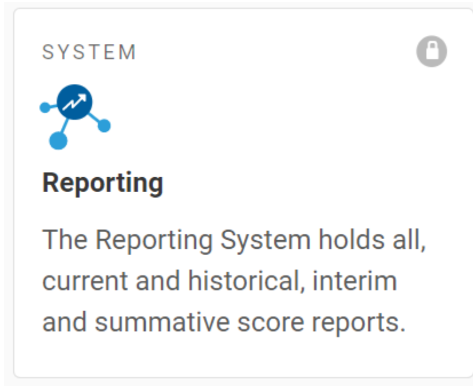
2. Select your assessment program from the cards displayed (see Figure 101).

Figure 101. Assessment Program Cards



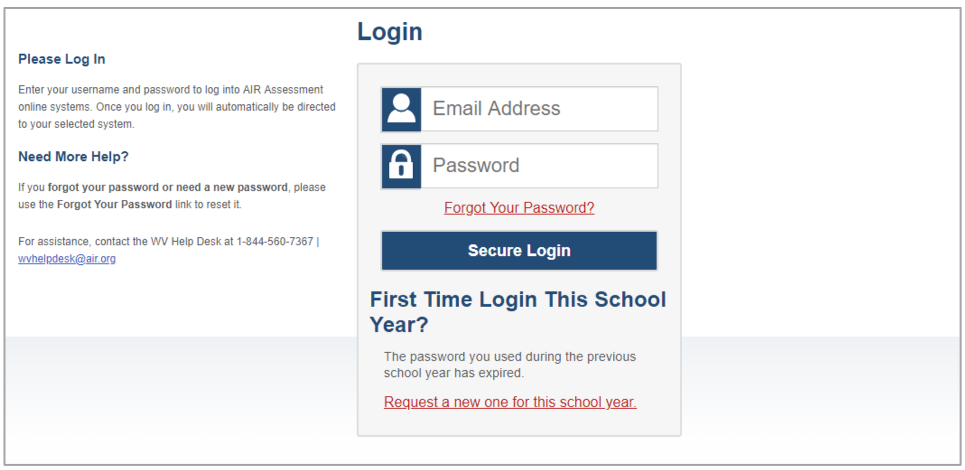
3. Click the **Reporting** card (see Figure 102). The login page appears.

Figure 102. Reporting Card



4. On the login page (see Figure 103), enter the email address and password you use to access all CAI systems.

Figure 103. Login Page

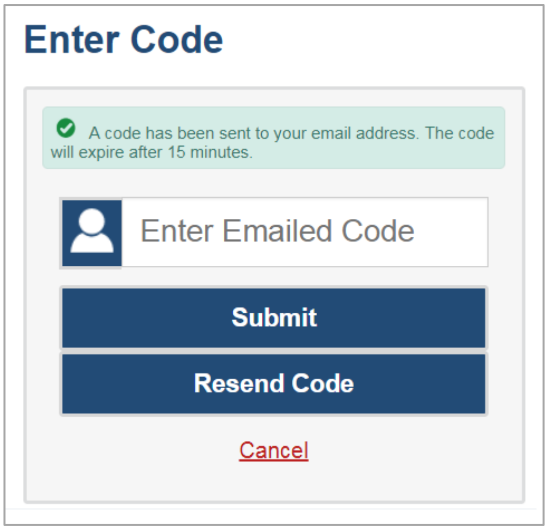


5. Click **Secure Login**.

   b. If the Enter Code page appears (see Figure 104), an authentication code is automatically sent to your email address. You must enter this code in the *Enter Emailed Code* field and click Submit within 15 minutes.

– If the authentication code has expired, click **Resend Code** to request a new code.

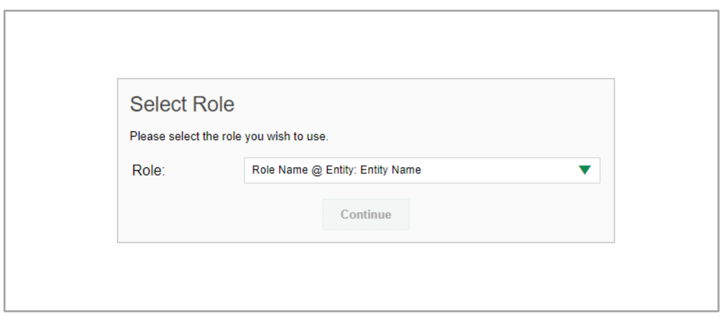
Figure 104. Enter Code Page



6. If the ***Terms and Conditions*** page appears, you should review the terms on this page and click **Accept** to proceed.

7. If your account is associated with multiple institutions, you are prompted to select a role, as in Figure 105. From the **Role** drop-down list, select the role and institution combination you wish to use. You can also change your institution after logging in.

Figure 105. Select Role Window



The Dashboard Generator for your user role appears.

How to Set or Reset Your Password

Your username is the email address associated with your account in TIDE. When you are added to TIDE, you receive an activation email containing a temporary link to the ***Reset Your Password*** page (see Figure 106). To activate your account, you must set your password within 15 minutes.

All users are required to do a one-time password reset at the beginning of every school year, for security purposes.

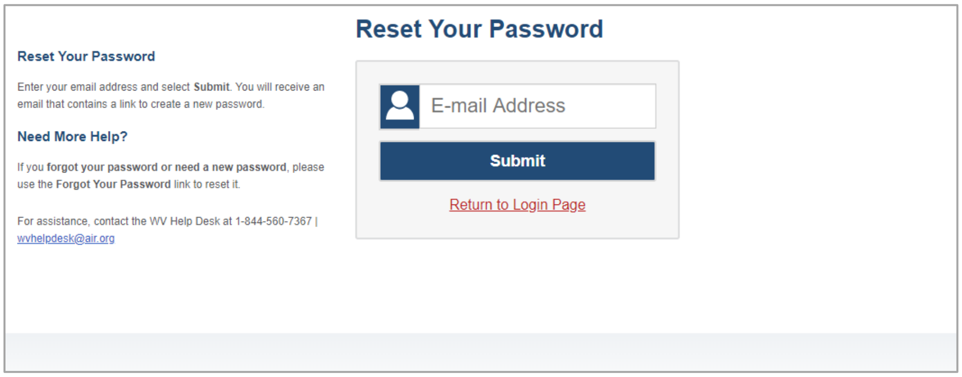
- If your first temporary link expired:

  In the activation email you received, click the second link provided and request a new temporary link.

- If you forgot your password:

  On the ***Login*** page, click **Forgot Your Password?** and enter your email address in the *E-mail Address* field. Click **Submit**. You will receive an email with a new temporary link to reset your password.

Figure 106. Reset Your Password Page



- If you did not receive an email containing a temporary link or authentication code:

  Check your spam folder to make sure your email program did not categorize it as junk mail. If you still do not see an email, contact your System Test Coordinator to make sure you are listed in TIDE.

- Additional help:

  If you are unable to log in, contact your Helpdesk for assistance. You must provide your name and email address.

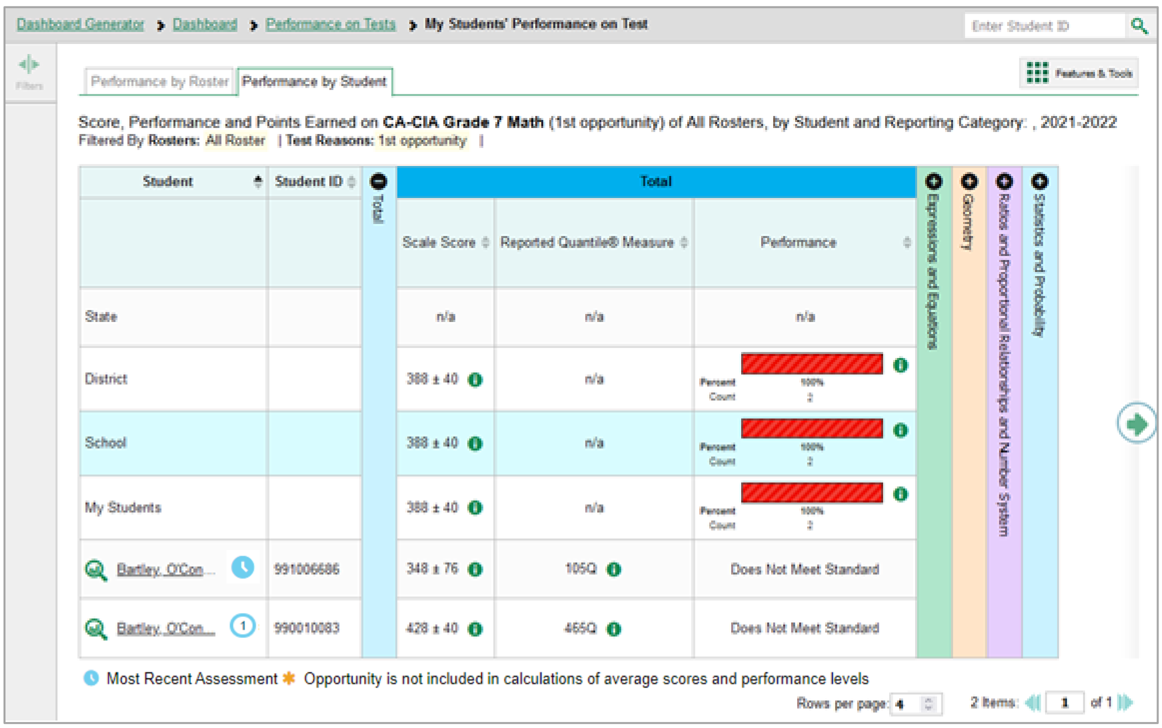
M

Multiple Interim Test Opportunities

Sometimes interim test results will include multiple rows for the same student.

When a student completes multiple test opportunities for a single assessment, as in Figure 107, reports display a row of data for each opportunity. A clock icon appears next to the most recent opportunity. Previous opportunities are marked with numbers, starting from the earliest test taken. An asterisk icon ✱ indicates that an opportunity is not included in calculations of average scores or performance distributions.

Figure 107. My Students' Performance on Test Report: Performance by Student Tab



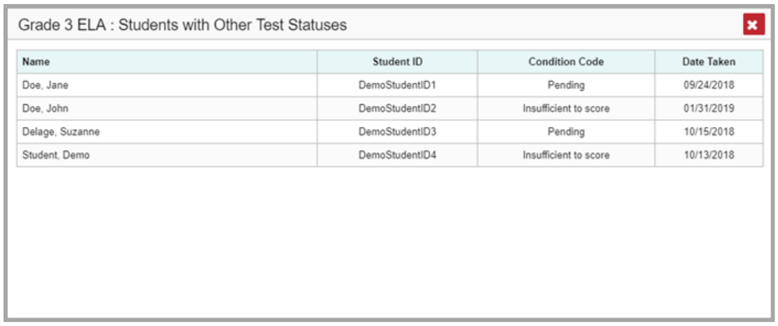
N

Non-Scorable Test Opportunities

The reports in Reporting do not include data for non-scorable test opportunities. A student's test opportunity cannot be scored when it has a test status of "Expired" or "Invalidated", or when it includes blank or empty reporting categories (reporting categories without items). If a test opportunity is non-scorable, a notification appears below the report for that assessment.

You can click **More Info** on the notification to view the ***Students with Other Test Statuses*** window (see Figure 108). This window lists the students who have non-scorable test opportunities for the given assessment, as well as the status code and completion date for each.

Figure 108. Students with Other Test Statuses Window


Grade 3 ELA : Students with Other Test Statuses

Name	Student ID	Condition Code	Date Taken
Doe, Jane	DemoStudentID1	Pending	09/24/2018
Doe, John	DemoStudentID2	Insufficient to score	01/31/2019
Delage, Suzanne	DemoStudentID3	Pending	10/15/2018
Student, Demo	DemoStudentID4	Insufficient to score	10/13/2018

P

Performance Data

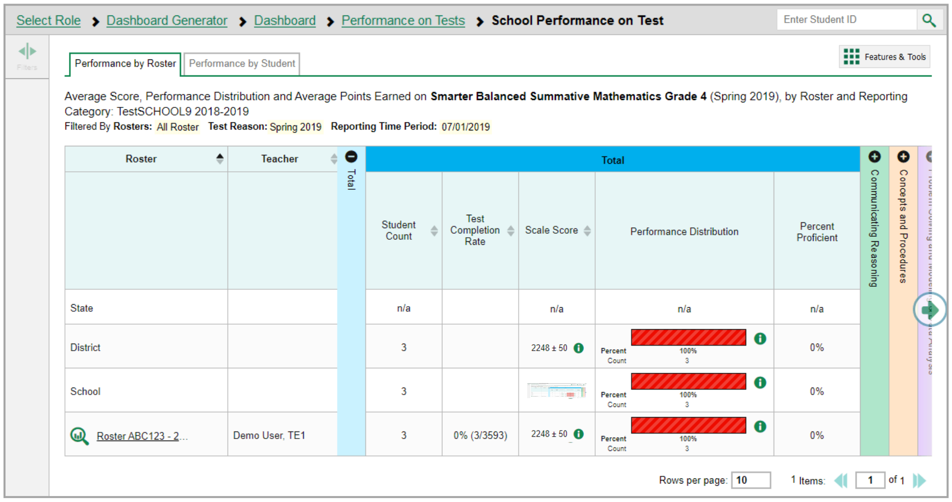
Depending on the test, a report may display different kinds of performance data:

- Score data:
  - Scale scores.
  - Raw scores, which may be in the form of percentages or fractions.
- Standard error: After an individual student's score, you may see a number with "±" before it. This is the standard error of measurement (SEM). A student's score is best interpreted when recognizing that the student's knowledge and skills fall within a score range and not just a precise number. As an example, if a student receives a test score of 75 with an SEM of 4, that tells us that the student's knowledge and skills fall between 71 and 79. For average scores at aggregate levels, the number following "±" is the standard error of the mean.
- Performance level data, which are used for tests with performance levels (also known as proficiency levels). Performance levels provide qualitative measurements of students' proficiency in relation to a particular standard or set of standards. Some aggregate reports include performance distribution bars, as in Figure 109, showing the percentage and number of students who achieved each performance level. These bars are color-coded, with three performance levels being coded red-

yellow-green, four being coded red-yellow-green-blue, and five being coded red-yellow-green-blue-purple.

- Percent proficient, also shown in some aggregate reports (see Figure 109), represents the total percentage of students who achieved proficiency. It typically includes those who fell into the top one to three performance levels.

Figure 109. School-Level User View: School Performance on Test Report: Performance by Roster Tab



- Measures in aggregate reports for adaptive tests may also include **Weak or Strong?**, **Proficient?**, and **% Correct**.

In a report, click the more information button in the score or Performance Distribution columns.

A legend appears (see Figure 110 and Figure 111), explaining what the scores or performance levels indicate.

Figure 110. My Students' Performance on Test Report with Expanded Scale Score Legend

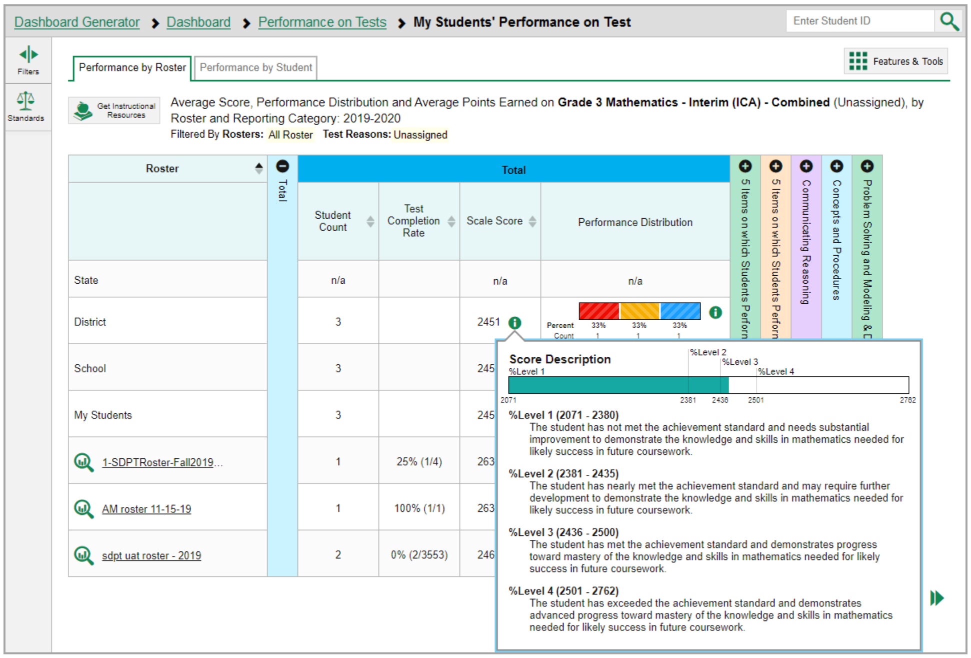
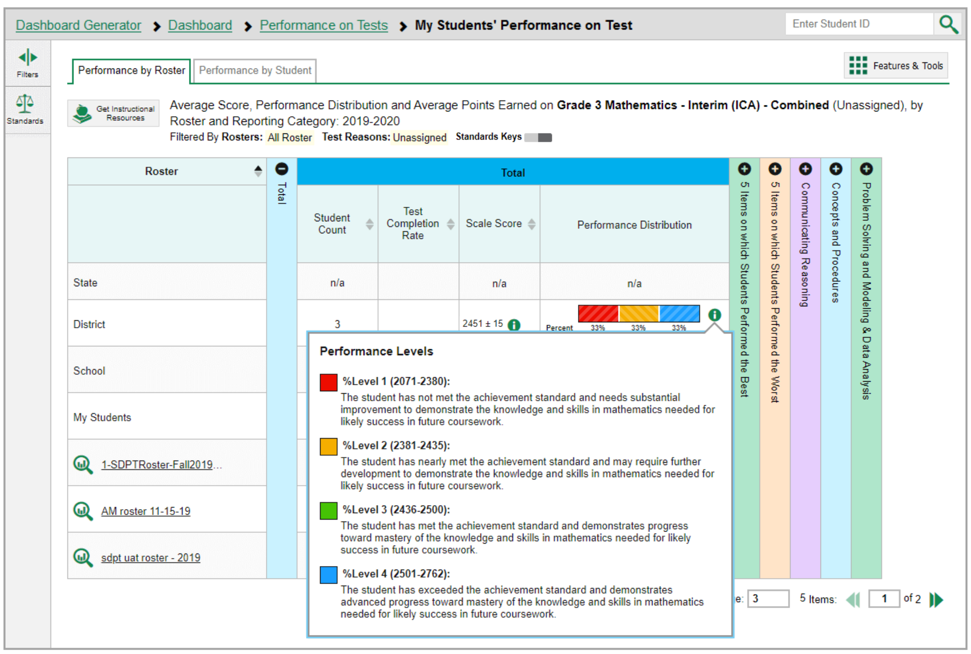


Figure 111. My Students' Performance on Test Report with Expanded Performance Distribution Legend



You will find similar buttons in reports throughout the Reporting System.

R

Report Tables

How to Sort a Table

1. To sort by descending order, click the header of the column you wish to sort by. The bottom arrow in the header is shaded darker when the column is sorted in descending order.

2. To sort by ascending order, click the column header again. The top arrow in the header is shaded darker when the column is sorted in ascending order.

How to Specify the Number of Rows Displayed

In the *Rows per page* field below a table, enter the number of rows you want the table to display per page. Your specifications persist for each table.

You can click the arrow buttons in this field to increase or decrease the number of rows displayed in increments of one.

How to View Additional Table Rows

- To move to the next and previous pages in a table, click the arrow buttons at the lower-right corner of the table.
- To jump to a specific table page, enter the page number in the field 1 at the lower-right corner of the table.

How to View Additional Table Columns

To scroll the table to the right or left, click the arrow buttons on the right and left sides of the table. Alternatively, click and drag the green horizontal scrollbar at the bottom of the table.

If a table contains expandable and collapsible accordion sections, you can click the section bars or and to expand and collapse them.

How to Expand All Accordion Sections in a Table

If you're navigating the page by tabbing through it, you may want to expand all the expandable accordion sections of a table at once. This feature, which is available in most test results, will make the table accessible to a screen reader.

1. Navigate to the table by tabbing through the page in your browser. When the “Load Accessible Table” message appears, press the **Enter** key. All the accordion sections expand.

2. *Optional*: To collapse the sections again, navigate back to the table. When the “Hide Accessible Table” message appears, press the **Enter** key. All the accordion sections collapse, except the **Total** section.

T

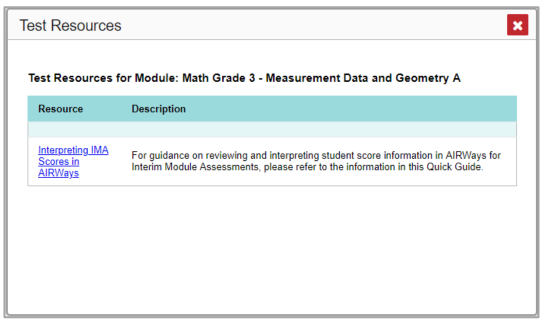
Test Resources

Some test results in the Reporting System include supplementary information that you can access, such as resources provided for the assessment in Tools for Teachers.

If additional assessment information is available, click **Get Instructional Resources** in the **Features & Tools menu**.

A window opens (see Figure 112), displaying resource links that either download or open in a new browser tab or window.

Figure 112. Test Resources Window



U

User Role Change

If your account is associated with multiple user roles or institutions, you can switch between them. The following instructions do not apply to switching schools for hand scoring purposes.

1. From the **Features & Tools** menu, select **Change Role**. (Alternatively, click the **Select Role** link in the path at the upper-left corner.) The ***Change Role*** window appears.
2. From the **Role** drop-down list, select the desired role and institution (entity name) combination.
3. Click **Continue**. The window closes and the Dashboard Generator page appears.

User Support

For additional information and assistance in using the Reporting System, contact the Montana Assessment Helpdesk.

The Helpdesk is open 6:00 AM - 6:00 PM MST (except holidays or as otherwise indicated on the Montana Testing portal).

Montana Help Desk

Toll-Free Phone Support: 877-365-7915

Email Support: mthelpdesk@cambiumassessment.com

- Please provide the Help Desk with a detailed description of your problem, as well as the following:
- If the issue pertains to a student, provide the associated district or school for that student. SSID and additional student PII can only be communicated via phone with the help desk**. Do not leave any student identifying information such as a student name, SSID, and/or personal characteristics in a voicemail or email.**
- If the issue pertains to a Test Information Distribution Engine (TIDE) user, provide the user's full name and email address.
- Any error messages and codes that appeared, if applicable.
- Affected test ID and question number, if applicable.
- Operating system and browser information, including version numbers (for example, Windows 11 (21H2) and Firefox 87 or macOS 12.3 and Safari 13).

Appendix 5-E
Test Administrator User Guide

Montana Office of Public Instruction

Online Testing System Test Administrator User Guide

2022–2023

Published August 24, 2022

Prepared by Cambium Assessment, Inc.

Cambium Assessment

Table of Contents

Introduction to the User Guide

This user guide supports Test Administrators (TAs) who manage testing for students participating in the Montana practice tests and operational interim and summative tests. The following user roles in TIDE have permission to create a test session in the Online Testing System and act as a Test Administrator: Authorized Representative (AR), System Test Coordinator (STC), Building Coordinator (BC), Test Administrator (TA), Proctor (PR). Refer to the User Roles and Access document for more information on user permissions.

Organization of the User Guide

The guide includes the following sections:

- How TAs Proctor Test Sessions in the TA Site
- How Students Sign in to the Student Testing Site and Complete Tests

There is also an alphabetized Appendix with additional information and instructions about the TA Site, the Secure Browser, and more.

Understanding the Online Testing System's Sites

The Online Testing System delivers Montana's online ELA, math, science, and alternate science tests and consists of practice testing sites and operational testing sites. The functionality of the practice sites is similar to that of the operational sites. However, the tests that are available in the practice and operational sites are different. Tests administered in the TA Training Site are for practice, whereas the tests provided in the TA Interface are operational and students' scores will be official.

- **Practice Sites**
  - **TA Training Site:** Allows TAs to practice administering tests.
  - **Student Training Site:** Allows students to practice taking tests online and using test tools. Students can log in to the testing site with their name and ID or as guests. They can either take proctored tests in sessions created by TAs in the TA Training Site or they can take non-proctored tests.
- **Operational Testing Sites**
  - **TA Interface:** Allows TAs to administer operational tests.
  - **Student Testing Site:** Allows students to take operational tests. Students must log in with their name and ID.

Throughout the rest of this user guide, "TA Site" refers to both the TA Interface and TA Training Site.

How TAs Proctor Test Sessions in the TA Site

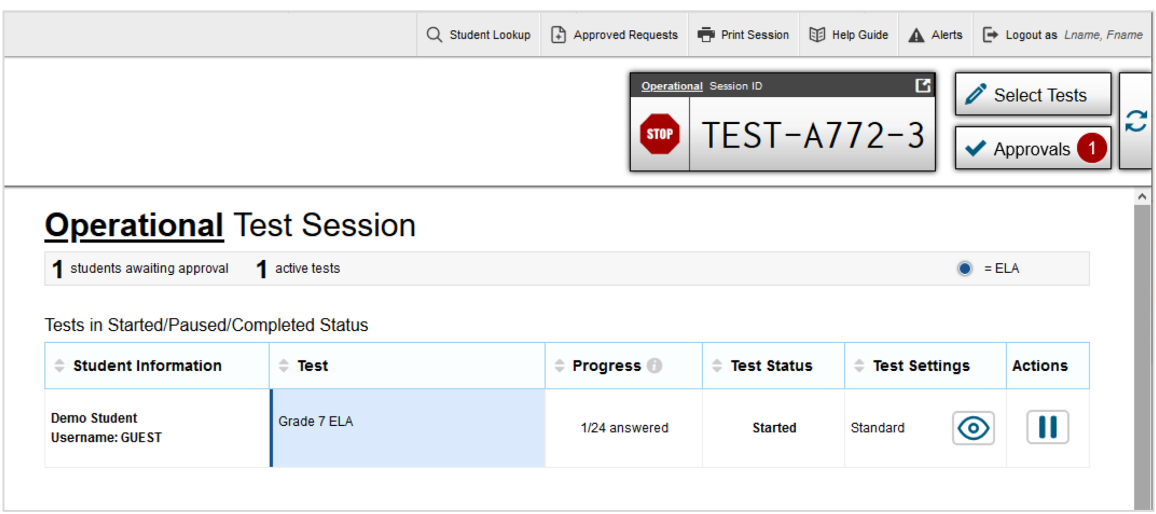
All users must complete the twenty-minute Test Administrator Certification module and acknowledge the virtual Test Security Agreement (TSA) when activating your TIDE account, prior to gaining access to any test related materials. Once these two steps are completed, administering tests in the Online Testing System is a straightforward process that follows the workflow below:

1. The TA selects tests and starts a test session in the TA Site (see Figure 1).
2. Students sign in to the Student Testing Site and request approval for tests.
3. The TA reviews students' requests and approves them for testing.
4. Students complete and submit their tests.
5. The TA stops the test session and logs out.

This section describes how TAs perform the following tasks within the TA Site to successfully administer online tests:

- How to Select Tests and Start a Test Session
- How to Approve Students for Testing
- How to Monitor an Ongoing Test Session

Figure 1. TA Site During an Ongoing Test Session



For information about the testing process from a student's perspective, see the section How Students Sign in to the Student Testing Site and Complete Tests.

How to Select Tests and Start a Test Session

The first steps for administering online tests are to select the tests that you wish to administer and start a test session.

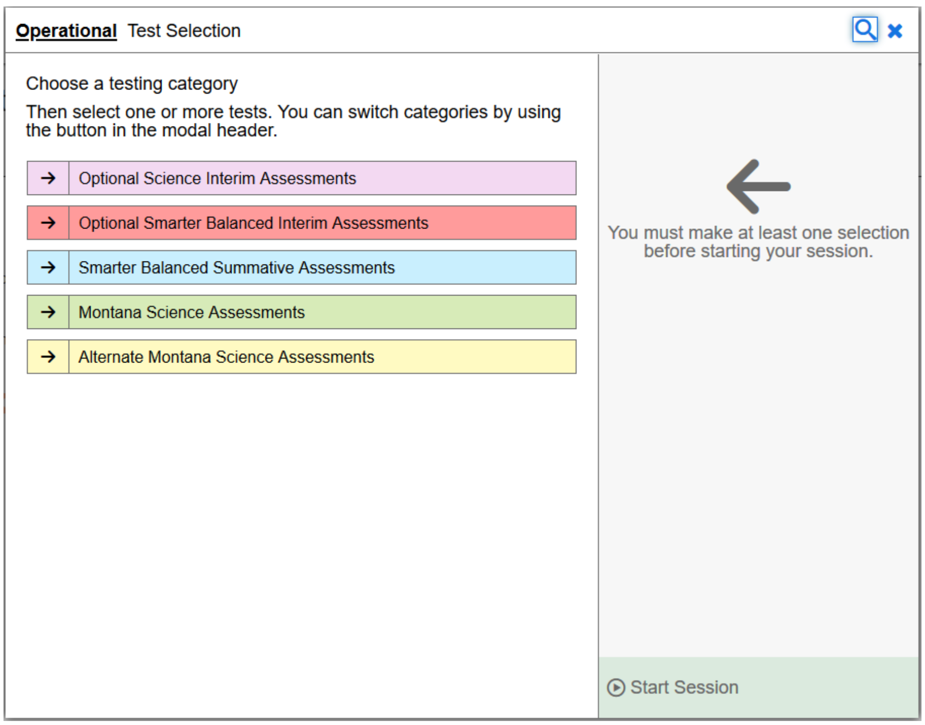
Please note that only the tests that you select will be available to students who join your session. You may have only one session open at a time. You cannot reopen closed sessions, but students can resume a test in a new session.

How to Create a New Test Session

1. Log in to the TA Site. The ***Test Selection*** window opens automatically (see Figure 2). If the ***Test Selection*** window is not open, click **Select Tests** in the top-right corner of the TA Site.

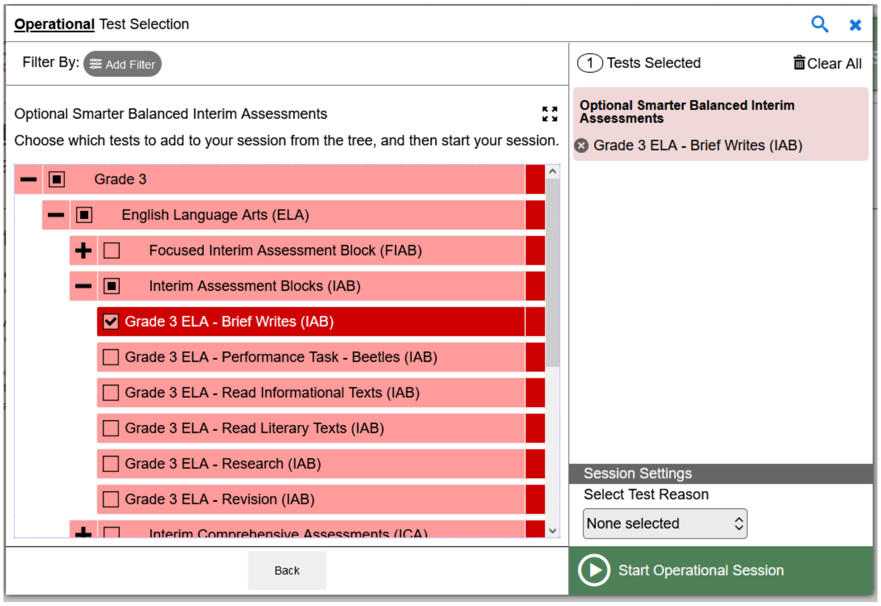
   – If you already have an active session running, you will need to select **Start a New Session Now** to open the ***Test Selection*** window. For more information, see Transfer a Test Session.

Figure 2. Test Selection Window: Test Categories



2. From the list of color-coded test categories, select the test category from which you wish to include tests. This displays the tests or test groups available for that category (see Figure 3).

Figure 3. Test Selection Window: Test Category Subgroups



3. *Optional*: To view the tests in a test group, select ✚ to expand it (or ⛶ to expand all groups).

   – To collapse an expanded test group, select ▬ (or to collapse all expanded groups).

4. *Optional*: If you wish to filter the available tests or search for specific tests, do either of the following. Otherwise, skip to step 5.

   – To filter tests by their subject or grade level, select Add Filter. In the filter panel that appears, expand the available filter categories and check the necessary grades or subjects that you wish to filter by. Select **Apply Filter(s)**. The test list updates to display the tests that match your filter criteria. If you need to remove a filter, select for that filter at the top of the window.

   – To search for a test, select in the top-right corner. In the *Search Term* field, enter a full or partial test label and select **Go**. The tests matching the entered label will appear. To close the search panel, select **Close** at the bottom of the panel.

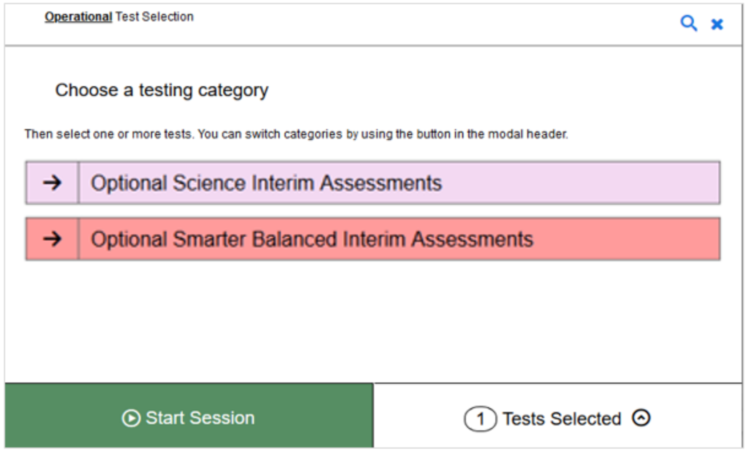
5. To select the tests you wish to administer, do one of the following:

   – To select individual tests, mark the checkbox for each test you want to include.

   – To select all the tests in a test group, mark the checkbox for that group.

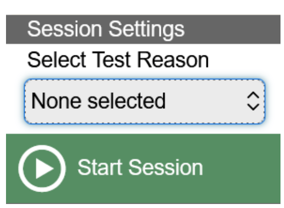
Once selected, tests display under their respective test categories in the panel on the right of the ***Test Selection*** window (see Figure 3). If viewing on a smaller screen, the test selection count displays at the bottom of the window (see Figure 4). To expand the selected tests section, select .

Figure 4. Test Selection Window: Tests Selected View for Small Screens



6. If a test alert message appears when selecting a test, review the message. Then, select **OK** to close the message and proceed. Alert messages may appear when:
   - Adding summative or high-stakes tests in a session.
   - Selecting a stand-alone test along with additional tests. A stand-alone test can only be administered by itself.
7. To add tests from a different test category, select **Back** at the bottom of the ***Test Selection*** window to return to the test categories view (see Figure 2). Then repeat steps 2–6 to select more tests.
8. *Optional*: If you need to remove selected tests, do one of the following:
   - To remove an individual test, select ⓧ for each test you want to remove.
   - To remove all the selected tests, select **Clear All**.
9. If any of the selected tests require you to select additional settings, such as a test reason, a drop-down list appears in the *Session Settings* section (see Figure 5). From the drop-down list, select the required setting for the session.
   - The **Test Reason** setting categorizes the test opportunities in your session for reporting purposes.

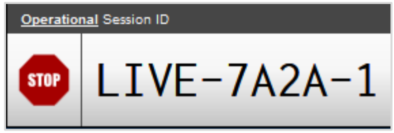
Figure 5. Session Settings Drop-Down List



10. Once the required tests have been selected, select **Start Session**. The exact label for this button may vary depending on whether you are starting a practice or operational session.

11. The Session ID appears on the TA Site (see Figure 6). Provide the Session ID to your students.

Figure 6. Test Session ID



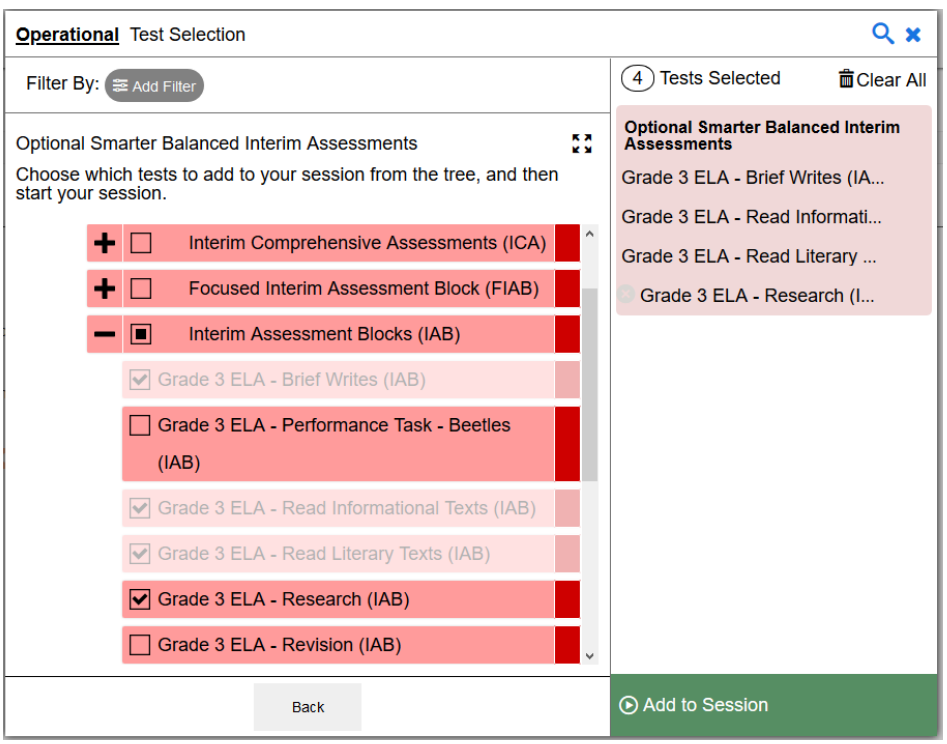
How to Add Tests to an Active Test Session

If necessary, you can add additional tests to an ongoing test session. While you can add tests to an active test session, you cannot remove tests from an active test session.

1. In the top-right corner of the TA Site (see Figure 1), click **Select Tests.** The ***Test Selection*** window (see Figure 7) opens to the page that you last viewed and shows the tests that are currently active in the session.

2. Mark the checkboxes of the tests that you wish to add to the session following the instructions in the How to Create a New Test Session section. Tests that are already active in the session or that cannot be added to the session are grayed out.

Figure 7. Test Selection Window: Add to Session



3. *Optional*: To remove a selected test:

   – To remove an individual test, select for each test you want to remove. The button displays only for tests that are not yet active in the session.

- To remove all the selected tests, select **Clear All**. All selected tests that are not yet active will be removed.

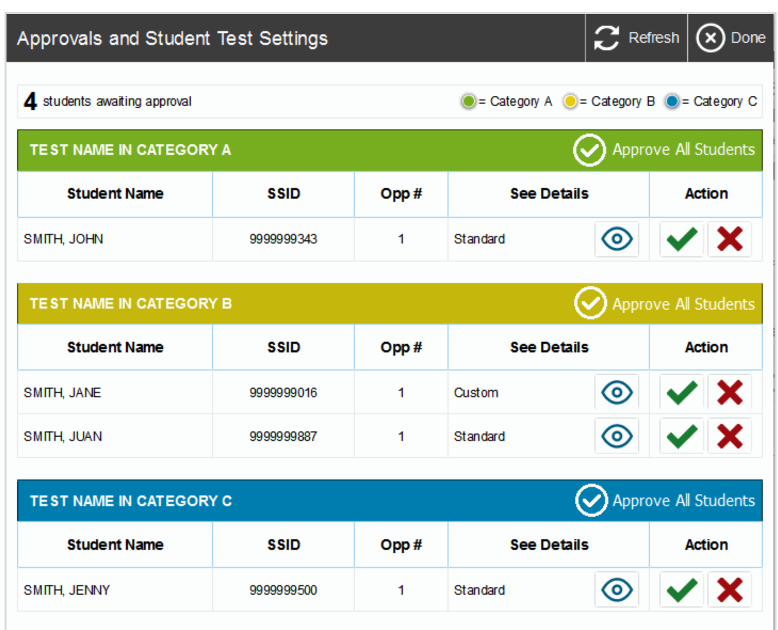
4. *Optional*: If any of the previously included or newly selected tests require session settings, a drop-down list appears in the *Session Settings* section (see Figure 5), displaying the test reason you selected when you started the session.
   - If you select a new test reason, then the test reason changes for every active test opportunity in the session. Any test opportunities that were completed before you changed the test reason will be submitted with the original test reason.
5. Select **Add to Session**. The exact label for this button may vary depending on whether you are starting a practice or operational session.
6. In the confirmation message that appears, click **Yes**.

How to Approve Students for Testing

After students sign in to the Student Testing Site and select tests, you must verify that their settings and accommodations are correct before approving them for testing. If a test uses segments that require TA approval, you must also follow the same procedure when approving students' entry to test segments.

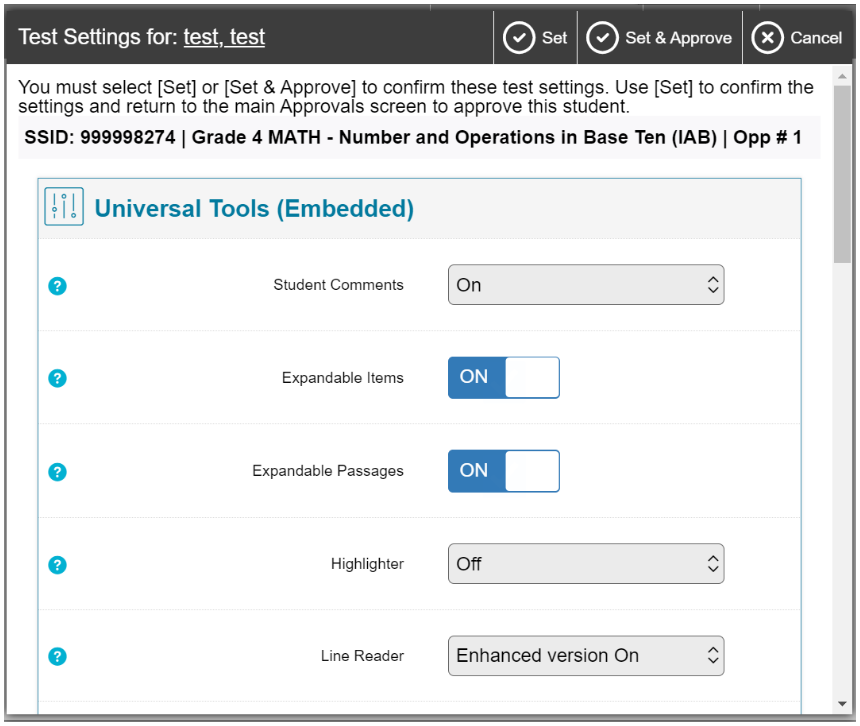
1. Once students request approval, select **Approvals** next to the Session ID (see Figure 1). The ***Approvals and Student Test Settings*** window appears (see Figure 8), displaying a list of students grouped by test. Note, the **Approvals** button becomes active when students are awaiting approval and shows you how many students are awaiting approval. The **Approvals** notification updates regularly, but you can also select in the upper-right corner to update it manually.

Figure 8. Approvals and Student Test Settings Window



2. To check a student's test settings and accommodations, select [eye icon] for that student. The ***Test Settings*** window appears (see Figure 9), displaying the student's test settings grouped by their area of need.

Figure 9. Test Settings Window for a Selected Student



a. If any settings are incorrect, update them as required. Note that accommodations can only be turned on by BC, or higher, user roles. Students should not begin testing until their settings are correct.

- Editable settings must be updated in this window, while changes to the read-only settings require changes in the district student data file system (Infinite Campus/PowerSchool). TIDE settings reflect the student data as it is uploaded to the state AIM database.

b. Do one of the following:

- To confirm the settings, select **Set**. You must still approve the student for testing (see step 5).
- To confirm the settings and approve the student, select **Set & Approve**.

3. Repeat step 2 for each student in the ***Approvals and Student Test Settings*** list. Since the **Approvals and Student Test Settings** window does not automatically refresh, select **Refresh** at the top of the window to update the list of students awaiting approval.

4. If you need to deny a student access to testing, do the following (otherwise skip to step 5):

a. Select [X icon] for that student.

b. *Optional:* In the window that appears, enter a brief reason for denying the student.

c. Select **Deny**. The student receives a message explaining the reason for the denial and is logged out. The student can still request access to the test again.

5. If you wish to approve students directly from the ***Approvals and Student Test Settings*** window, do the following:

– To approve individual students, select ✔ for each student.

– To approve all students for a given test or segment, select **Approve All Students** for that test or segment.

How to Monitor an Ongoing Test Session

After you approve students for testing, you can monitor the testing progress for each student logged in to your session, approve a student's print request, and pause a student's test, if necessary.

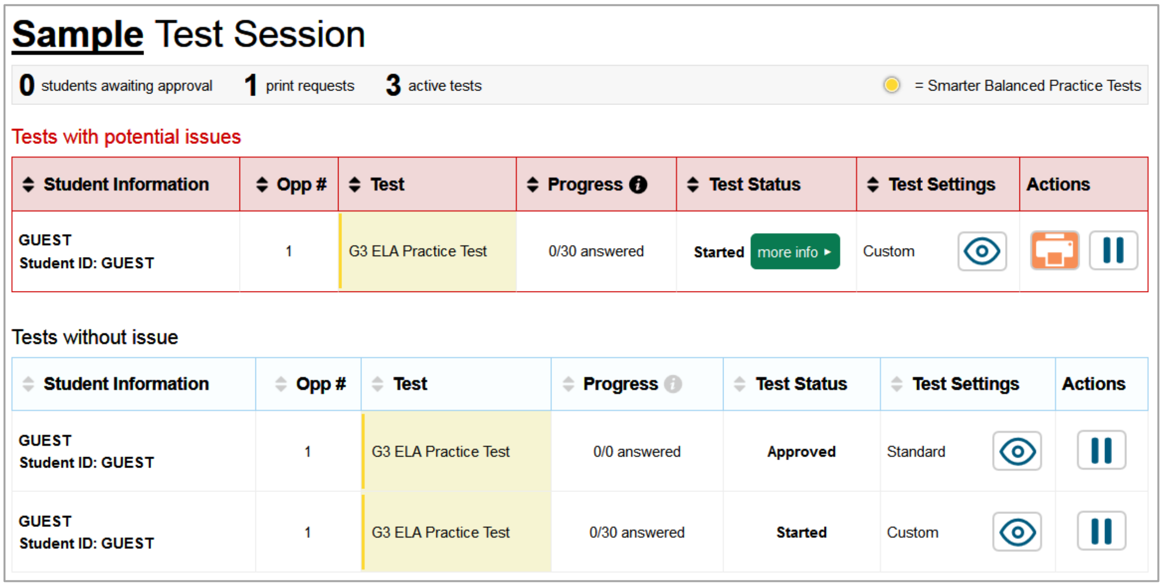
How to Monitor Students' Test Progress

You can monitor the testing progress for each student logged in to your session from the tables displayed on the TA Site.

At the start of the test, all the students in the session are listed in the Tests Without Issues table. If the Online Testing System detects that a student requires assistance, the Tests With Potential Issues table appears at the top, listing the students who need intervention. This can occur when a student has a pending print request or a student's test has been paused due to a security issue, such as the launch of a forbidden application.

The tables refresh at regular intervals, but you can also refresh the tables manually by selecting in the top-right corner of the TA Site. You can also sort the tables by a given column by selecting in that column's header.

Figure 10. Tables for Monitoring Students’ Test Progress



Sample Test Session

0 students awaiting approval **1** print requests **3** active tests = Smarter Balanced Practice Tests

Tests with potential issues

Student Information	Opp #	Test	Progress	Test Status	Test Settings	Actions
GUEST Student ID: GUEST	1	G3 ELA Practice Test	0/30 answered	Started more info ▸	Custom	

Tests without issue

Student Information	Opp #	Test	Progress	Test Status	Test Settings	Actions
GUEST Student ID: GUEST	1	G3 ELA Practice Test	0/0 answered	Approved	Standard	
GUEST Student ID: GUEST	1	G3 ELA Practice Test	0/30 answered	Started	Custom	

Table 1 describes the columns in the tables for monitoring students’ test progress.

Table 1. Columns in the Tables for Monitoring Students’ Test Progress

Column	Description
Student Information	The name and SSID of the student in the session.
Opp #	Opportunity number for the student’s selected test.
Test	Name of the test the student selected. For segmented tests, this column also displays the name of the test segment that the student is currently in.
Progress	Indicates the student’s test progress. It displays a progress bar to indicate how far the student has progressed in the test. The progress bar indicates the percentage of questions the student has answered out of the total number of questions, the percentage of questions the student has answered or viewed out of the total number of questions, or the percentage of questions the student has answered and the percentage of questions the student has viewed out of the total number of questions.
Test Status	Current status for each student in the session. For more information about the statuses in this column, see Table 2. If the Online Testing System detects that a student may be experiencing technical difficulties or requires assistance, such as the student may be experiencing connection issues, has a pending print request, or has paused his test, a more info icon (more info ▸) displays in this column for the student. When you hover over the icon, a message provides details about the issue.
Test Settings	Displays one of the following: • **Standard:** Default test settings are applied for this test opportunity. • **Custom:** One or more of the student’s test settings or accommodations differ from the default settings.

Column	Description
	To view the student's settings for the current test opportunity, select .
Actions	Allows you to perform any available actions for an individual student's test. The button in this column allows you to pause the student's test. If a student pauses their test, a more info icon (more info ▸)in the Test Status column provides information about how the test became paused. However, the more info icon does not display when the TA pauses a student's test. A button appears in this column when the student requests a printout of test material. For information, see the section How to Approve a Student's Print Request.

Table 2 describes the codes in the Test Status column of the tables for monitoring students' test progress.

Table 2. Student Testing Statuses

Column	Description
Approved	You approved the student, but the student did not yet start or resume the test.
Started	Student started the test and is actively testing.
Review	Student visited all questions and is currently reviewing answers before completing the test.
Completed	Student submitted the test. The student can take no additional action at this point.
Submitted	Test was submitted for quality assurance review and validation.
Reported	Test passed quality assurance and is undergoing further processing.
Paused*	Student's test is paused. The time listed indicates how long the test has been paused.
Expired*	Test was not completed by the end of the testing window and the opportunity expired.
Pending*	Student is awaiting approval for a new test opportunity.
Suspended*	Student is awaiting approval to resume a test opportunity.

*Appears when the student is not actively testing. The student's row grays out in such cases.

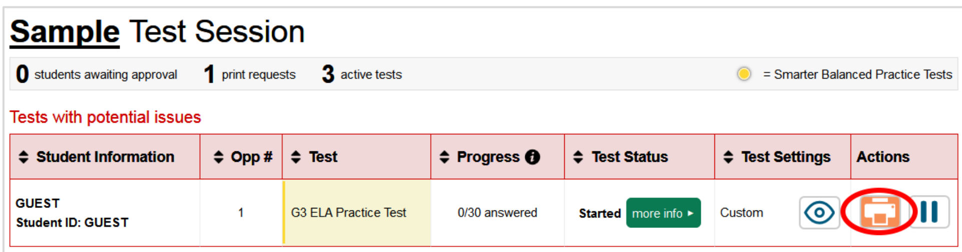
How to Approve a Student's Print Request

Students using the print-on-request tool can request printouts of test passages and questions. You must review and approve these print requests. When students send print requests, the request notification appears in the **Tests With Potential Issues** table.

You can also view a list of every print request you approved during the current session. For more information, see Print Approved Requests Information.

1. To review a print request, select [print icon] in the Actions column of the Tests With Potential Issues table for a student (see Figure 11). The request notification appears for students who have sent print requests.

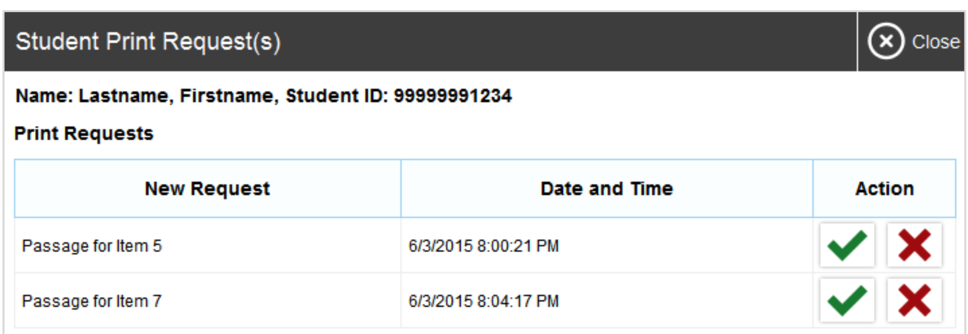
Figure 11. Print Request Notification



2. Review the request in the ***Student Print Request*** window and do one of the following:

   – To approve the request, select [check mark]. A cover sheet appears in a new browser window.

Figure 12. Student Print Request Window



   – To deny the request, select [X]. In the window that appears, enter a brief reason for denying the request and select **Deny**. Do not proceed to the next step.

3. In the window displaying the cover sheet, select **Print** to open the printer dialog box.

4. Select **OK** to print the requested test elements.

How to Pause a Student's Test

You can pause a student's test if necessary. Please note that pausing a timed test pauses the timer for that individual student but not the timer for the test session.

1. In the Actions column of the tables for monitoring students' test progress, select for the student whose test you wish to pause.

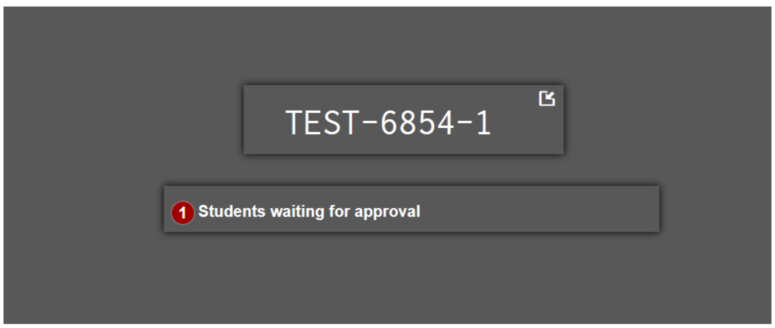
2. Select **Yes** to confirm. The Online Testing System logs out the student.

How to Enable Screensaver Mode

Since the student test progress tables in the TA Site often contain sensitive student information, the TA Site provides a built-in screensaver to hide the data from view. If the screensaver mode is auto-enabled, the screensaver will automatically turn on if you are not active in the TA Site for 5 minutes. If the screensaver mode is not auto-enabled, it is strongly recommended that you manually turn on screensaver mode when stepping away from your device.

1. To turn on screensaver mode, select in the top-right corner of the Session ID (see Figure 1). A masking screen appears over the TA Site (see Figure 13). The screensaver displays the Session ID and the timer, if applicable. It also displays notifications if students are awaiting approval, if there are pending print requests, or if students require other interventions.

Figure 13. Screensaver Mode Enabled



2. To turn off the screensaver mode, click . It may also deactivate when mouse or keyboard activity is detected, if such a feature is enabled.

How to Stop a Test Session

When students finish testing, or the current testing timeslot is over, you should stop the test session. Stopping a session automatically logs out all the students in the session and pauses their tests.

Once you stop a test session, you cannot resume it. To resume testing, you must start a new session. Please note, the Online Testing System automatically logs you out after 20 minutes of both user and student inactivity. This action automatically stops the test session.

1. To stop a test session, select STOP next to the Session ID (see Figure 1).

2. In the confirmation message that appears, select **OK**. The test session stops.

How to Log Out of the TA Site

To avoid stopping a test session that is in progress, you should only log out of the TA Site after stopping a test session. Please note that navigating away from the TA Site also logs you out. If you need to access another application while administering tests, open it in a separate browser window.

1. To log out, select in the top-right corner of the TA Site (see Figure 1). A warning message appears.

2. In the warning message, select **Log Out**. The State Assessment Program portal appears.

How Students Sign in to the Student Testing Site and Complete Tests

This section describes the student sign-in process for the Student Testing Site that students follow when starting a new test or resuming a paused test. It also describes how students can view stimuli, respond to questions, pause a test, review previously answered questions, and submit a test.

How Students Sign in and Select Tests

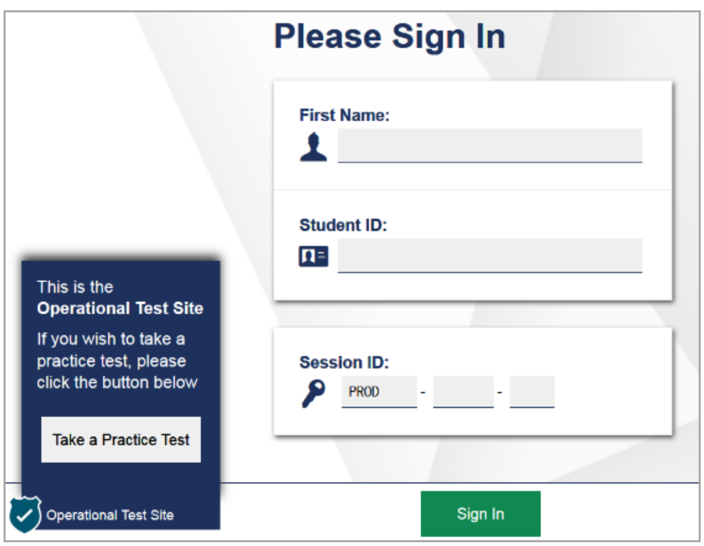
When testing, students must sign in to the appropriate testing site. For sessions created in the TA Interface, students sign in to the Student Testing Site on the Secure Browser.

Students may also take practice tests in the Student Training Site to familiarize themselves with the online testing process. Aside from the sign-in process, the Student Training Site has the same appearance and functionality as the Student Testing Site. For information on how students sign in to the Student Training Site, see Practice Test Site Student Sign-in Process.

How to Sign in to the Secure Browser

1. Launch the Secure Browser on the student's testing device. The ***Student Sign-In*** page appears.
2. Next, students enter the following information:
   a. In the *First Name* and *SSID* fields, students enter their first name and SSID as they appear in TIDE.
   b. In the *Session ID* field, students enter the Session ID as it appears on the TA Site. The first part of the session ID, which indicates whether a student is on the Student Testing Site or the Student Training Site, is pre-filled.

Figure 14. Student Testing Site Student Sign-In Page



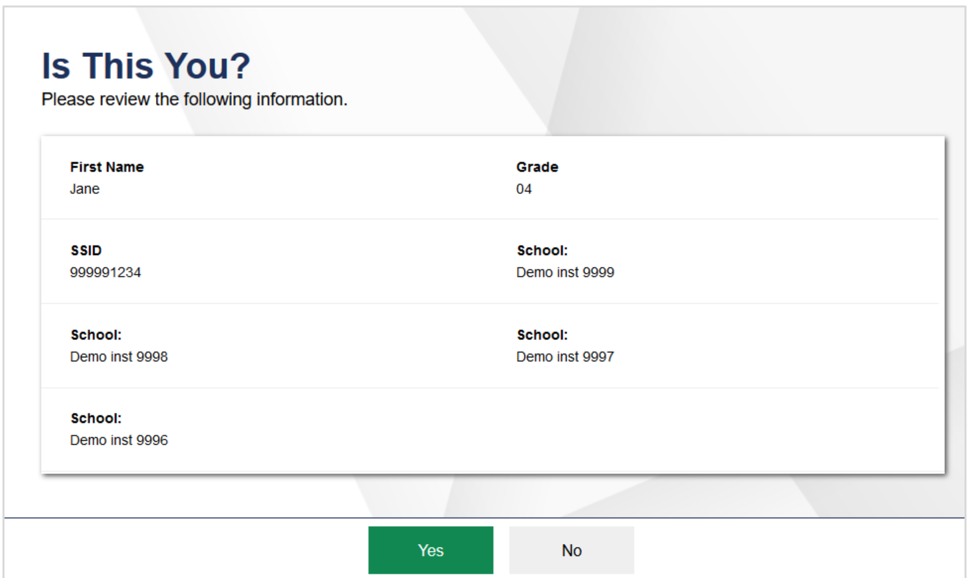
3. Students select **Sign In**. The ***Is This You?*** page appears.

How to Verify Student Information

After signing in to the Student Testing Site, students must verify their personal information on the ***Is This You?*** page.

- If all the information on the ***Is This You?*** page is correct, the student selects **Yes** to proceed.
- If any of the information is incorrect, the student must select **No**. You must notify the appropriate school personnel that the student's information is incorrect. Incorrect student demographic information must be updated before the student begins testing

Figure 15. Is This You? Page



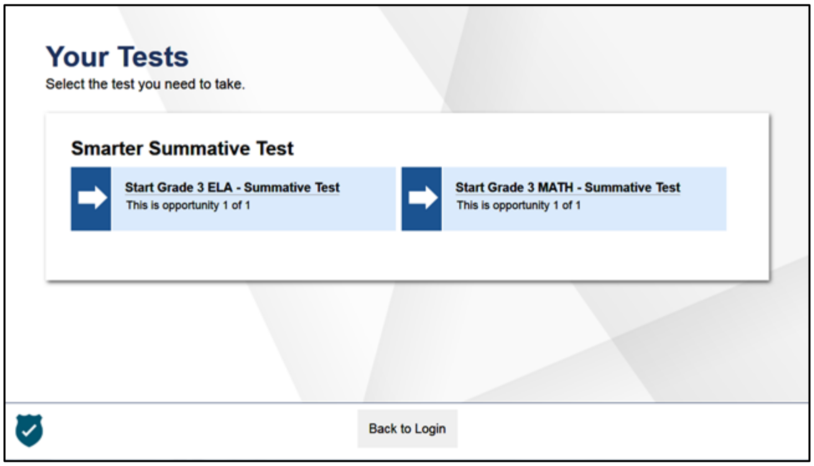
How to Select a Test

Students can select their tests from the ***Your Tests*** page that appears after students verify their personal information. The ***Your Tests*** page displays all the tests that a student is eligible to take, color-coded by test category. Students can only select tests that are included in the session and still need to be completed.

1. From the ***Your Tests Page***, the student selects the name of the test.

   – If a student's required test is inactive or not displayed, the student should log out. You should verify the test session includes the correct tests.

Figure 16. Your Tests Page

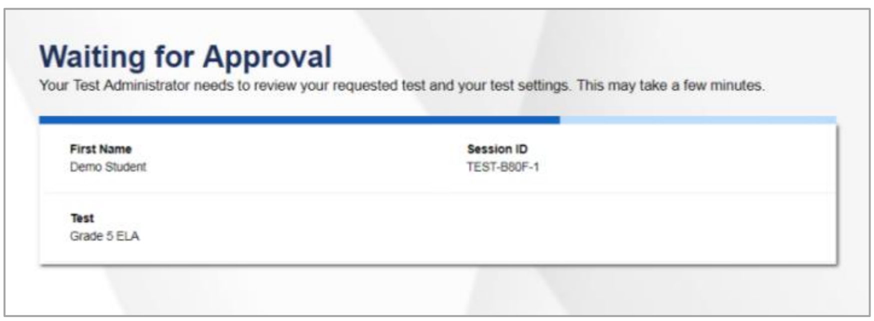


2. The student's request is sent to the TA for approval and the student is taken to the ***Waiting for Approval*** page. After you approve the student for testing, the student can proceed to the next step.

   – If starting a new test, the student must complete the login process before they begin testing.

   – If resuming a paused test, the student is taken directly to the test page where they stopped the test, based on the applicable pause rules.

Figure 17. Waiting for Approval Page



Please note that once the TA approves the test, the test content displays to the student in the language specified for the student's test.

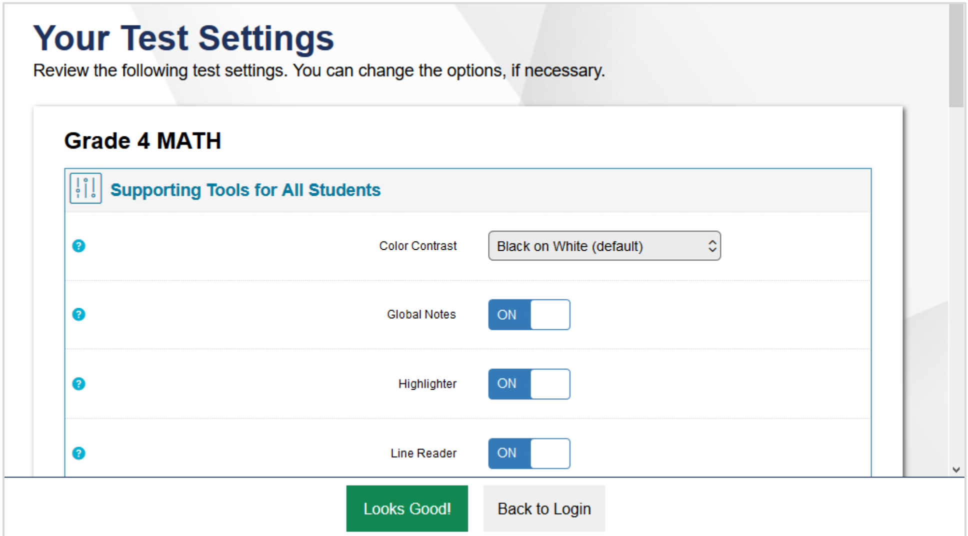
How to Verify Test Setting Information

Once students have been approved for testing, they can verify their test settings from the ***Your Test Settings*** page. At this point, the student's actual test settings override any settings selected earlier in the sign-in process.

- If the settings are correct, the student selects **Looks Good** to continue.
- If the settings are incorrect, the student should select **Back to Login** to log out of the Student Testing Site.

After you correct the student's test settings, the student must sign in and request approval again.

Figure 18. Your Test Settings Page



How to Check Student Device Functionality

Depending on the test content and the specified test settings, students may need to verify that their testing device is functioning properly before starting the test. If a test does not require functionality checks, this page is skipped.

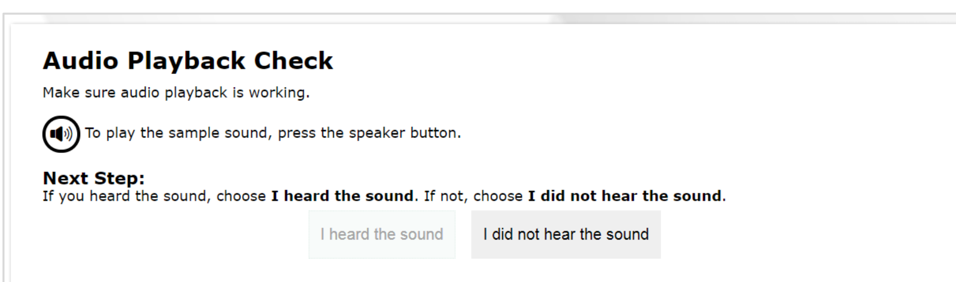
How to Check Audio Playback Functionality

The ***Audio Playback Check*** panel appears for tests with listening questions and allows students to verify that they can hear the sample audio.

If the audio does not work, students should log out. You should troubleshoot the device and headphones or move the student to another device with working audio.

- From the ***Audio Playback Check*** panel, students select and listen to the audio.
  - If the sound is clearly audible, students select **I heard the sound**. A green check appears at the top-right corner of the panel and students can proceed to the next functionality check.

Figure 19. Audio Playback Check Panel



  - If the sound is not clearly audible, students select **I did not hear the sound** to open the ***Sound Check: Audio Problem*** panel.

- Students can select **Try Again** to return to the ***Audio Playback Check*** panel and retry.

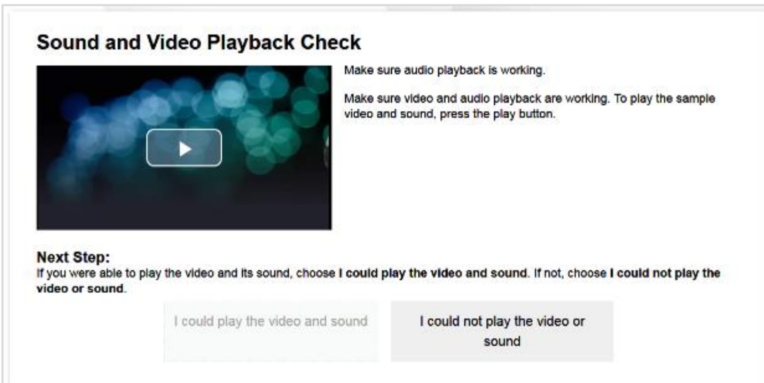
How to Check Sound and Video Playback Functionality

The ***Sound and Video Playback Check*** panel appears for tests with video content and allows students to verify that they can view the sample video and hear its associated sound.

If the video or audio does not work, students should log out. You should troubleshoot the device and headphones or move the student to another device with working audio and video.

- From the ***Sound and Video Playback Check*** panel, students select ▶ to play the video and sound.
  - If the video plays and the sound is clearly audible, students select **I could play the video and sound**. A green check appears at the top-right corner of the panel and students can proceed to the next functionality check.

Figure 20. Sound and Video Playback Check Panel



  - If students are not able to play the video or hear the sound, students select **I could not play the video or sound** to open the ***Video Playback Problem*** panel.
    - Students can select **Try Again** to return to the ***Sound and Video Playback Check*** panel.

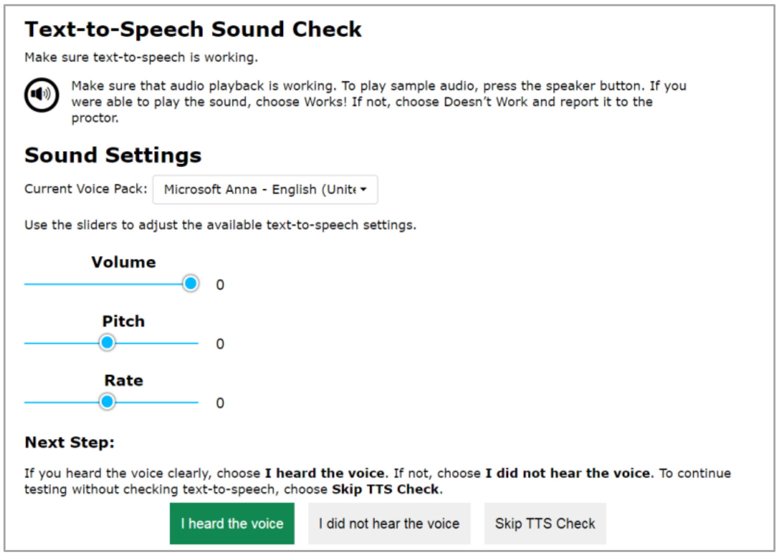
How to Check Text-to-Speech Functionality

The ***Text-to-Speech Sound Check*** panel appears if a student has the Text-to-Speech (TTS) setting.

If TTS does not work, students should log out. You can work with students to adjust their audio or headset settings or move them to another device.

- From the ***Text-to-Speech Sound Check*** panel, students select 🔊 and listen to the audio.
  - If the voice is clearly audible, students select **I heard the voice**. A green check appears at the top-right corner of the panel and students can proceed to the next functionality check.
  - If the voice is not clearly audible, students adjust the settings using the sliders and select 🔊 to listen to the audio again.

Figure 21. Text-to-Speech Sound Check Panel



- If students still cannot hear the voice clearly, they select **I did not hear the voice** to open the ***Audio Check*** panel.
  - Students can select **Try Again** to return to the ***Text-to-Speech Sound Check*** panel and retry.
  - Students can select **Continue** to skip verifying the TTS functionality. Students can also do this from the ***Text-to-Speech Sound Check*** panel by selecting **Skip TTS Check**.

How to Check Recording Device Functionality

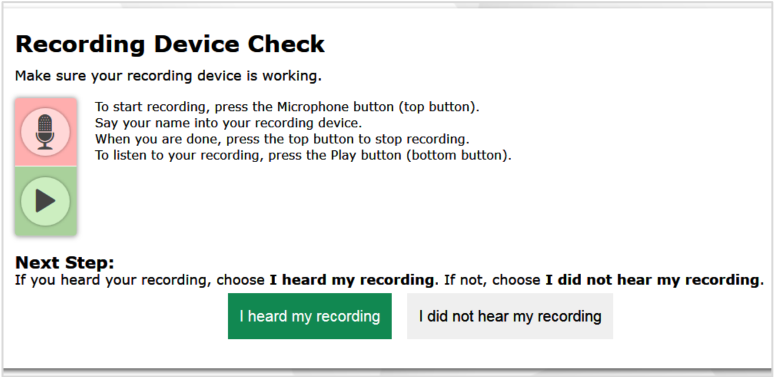
The ***Recording Device Check*** panel displays for students who have a Speech-to-Text (STT) accommodation. On this panel, students record their voice and verify that they can hear the recorded audio.

If the recording device does not work, students should log out. You should troubleshoot the recording device or set up a new recording device.

3. From the ***Recording Device Check*** panel, students select  to begin recording their voice.

4. Students speak into their recording device, and then select  to stop recording.

Figure 22. Recording Device Check Panel



5. Students select ▶ to listen to their recorded audio:

   – If the recorded audio is clearly audible, students select **I heard my recording**. A green check appears at the top-right corner of the panel and students can proceed to the next functionality check.

   – If the recorded audio is not clearly audible, they select **I did not hear my recording** to open the ***Problem Recording Audio*** panel.

     - Students can select **Try Again** to return to the ***Recording Device Check*** panel and retry.

     - If the student is testing on a computer or tablet with multiple recording devices, they can select the **Select New Recording Device** option to open the ***Recording Input Device Selection*** panel and select a different recording device.

How to View Instructions and Begin Testing

The ***Instructions and Help*** page is the last step of the sign-in process. Students may review this page to understand how to navigate the test and use test tools as well as review their test settings. This page may also contain additional test instructions or acknowledgements that students need to review in order to proceed.

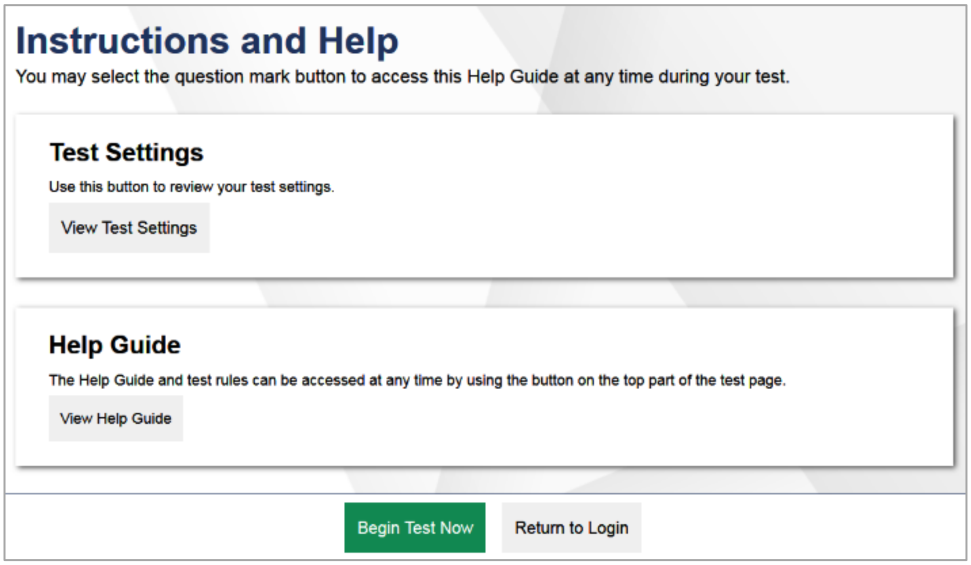
1. *Optional*: To review their test settings, students select **View Test Settings**. To close the window, students select **OK**.

2. To review additional test information or acknowledgements, if available, students review the information on the page and proceed as instructed.

3. *Optional*: To view the help guide, students select **View Help Guide**. To close the window, students select **Back**.

4. To start the test, students select **Begin Test Now**.

Figure 23. Instructions and Help Page



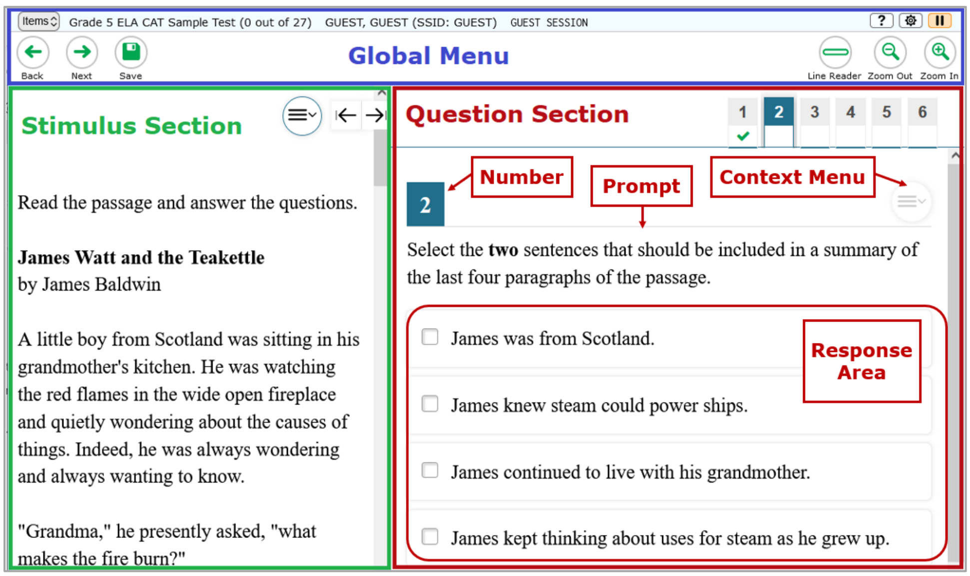
How Students Navigate the Student Testing Site

This section explains how students use the features available in the Student Testing Site to navigate tests, use tools, and respond to items.

A test page in the Student Testing Site can include the following sections:

- The *Global Menu* section displays the global navigation and tool buttons. It also includes the **Items** menu, test information, help button, pause button, system settings button, and timer (if available).
- The *Stimulus* section appears only for test pages with a stimulus. This section contains the stimulus content, context menu, and either the expand passage button or reading mode button.
- The *Question* section contains one or more test questions (also known as "items"). Each question includes a number, context menu, prompt, and response area. Each question also displays the student's name and the question's most recent save date and time.

Figure 24: Sample Test Layout



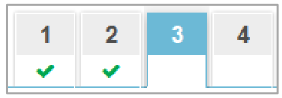
How to Navigate Between Items

- Some test pages may have only one question and others may have more.
  - After students respond to all the questions on a page, they select **Next** in the top-left corner to proceed to the next page.
  - To navigate to a previous question in a test, students select **Back**.

Figure 25. Navigation Buttons

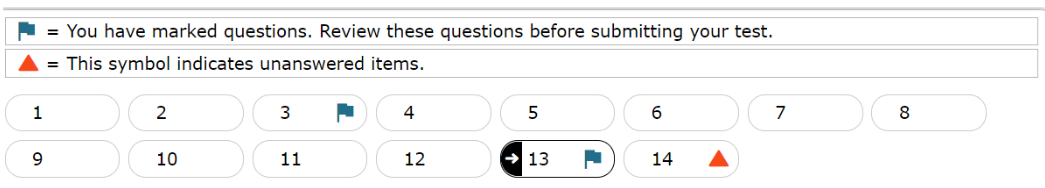
- When multiple items are grouped with a stimulus, the items are tabbed for individual viewing. Students select the tabs in the top-right corner to proceed through the questions.

Figure 26. Paginated Items



- To jump directly to an item, students can select an item number from the **Items** menu in the top-left corner.
  - If an item has been marked for review, [flag icon] displays next to the item.
  - If an item has been skipped or left unanswered, [triangle icon] displays next to the item.
  - Items that students cannot navigate to are grayed out.

Figure 27. Items Drop-Down



How to View Stimuli

When a test question is associated with a stimulus, students should review that stimulus before responding to the question. A stimulus is a reading passage or other testing material (such as a video or graphic) that students review in order to answer associated questions.

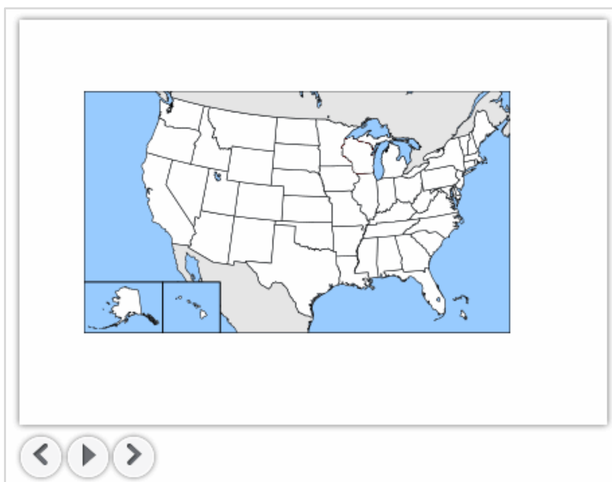
- *Videos*: When the stimulus is a video, students can use standard video features to control the playback.

Figure 28. Video Playback Features

- To play a video, select in the lower-left corner.
- To jump to a different point in the video, drag the slider to the required location.
- To adjust the speed at which the video plays, select 1x, and then select the required speed from the menu that appears.
- To mute or unmute the video, select in the lower-right corner.
- To expand the video to full-screen mode, select in the lower-right corner. To exit full-screen mode, select again.

- *Slideshows*: When the stimulus is a slideshow, students can navigate between the slides and play associated audio, if available.
  - To move between the slides of a slideshow, select and .
  - To play the audio for the current slide, select . To pause the audio, select .

Figure 29. Sample Slideshow



How to Respond to Test Questions

The questions available in TDS may use various item types that require students to respond to them in different ways. Students can use the Student Training Site to familiarize themselves with the question types that may appear on tests.

All responses are saved automatically. Students can also manually save their responses to questions by selecting **Save** in the top-left corner.

Test questions may require students to do any of the following tasks:

- Select one or more choices from a list of answer options.
  - For multiple choice type items, students can re-click a selected radio button to deselect the response option, provided this feature is enabled.
- Use an on-screen keypad to generate an answer. Students can select in the answer space to open the keypad.
- Select graphic objects or text excerpts.
- Place points, lines, or bars on a graph.
- Drag and drop text or graphic objects.
- Enter text in a text box or table. For more information, see the section Text-Response Questions.
- Match answer options together.
- Modify a highlighted word or phrase in a reading selection.
- Enter input parameters to run an on-screen simulation.
- Copy content from a passage to a text box.
- Expand categories and select options within them.
- Create graphs and charts.
- Record a spoken response using a microphone. For more information, see the section Recording Spoken Responses.

How to Pause Tests

Students can pause the test at any time. Pausing a test logs out the student. To resume testing, students must repeat the sign-in process.

- To pause a test, students select **Pause** in the global menu and then select **Yes** in the confirmation message that appears.

If students are testing on Chromebooks, please ensure that they pause the test before closing the lid of the Chromebook. If the lid is closed before the test pauses, whoever opens the Chromebook next will be able to see the last question that the student viewed (and any response they entered).

How Students Use Test Tools

A number of testing tools are available for students in TDS. Some tools are available for all tests, while others are available only for a particular subject, accommodation, or type of question. There are primarily two types of test tools available:

- **Global Tools**: These tools appear in the global menu at the top of the test page and are available for all items in a test.
- **Context Menu Tools**: These tools are specific to the passage or question being viewed.

Students can access tools using a mouse or keyboard commands. For information about keyboard commands, please see Keyboard Commands in the Student Testing Site.

Some tools open in separate windows. Students can resize the window using the resizing handle. Students can also maximize or minimize the window by selecting. Students can also drag and move the window to different parts of the test page.

How to Use Global Tools

The global menu consists of navigation buttons on the left and tool buttons on the right (see Figure 30).

- To use a global test tool, select the button for the tool. The selected test tool activates.

Table 3 lists the tools available in the global menu.

Figure 30. Global Menu

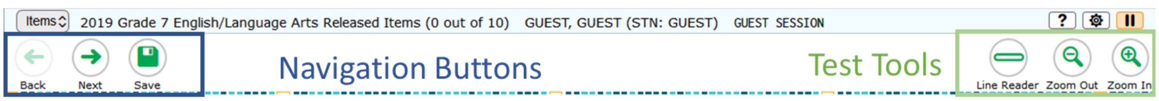


Table 3. Global Tools

Tool Name	Instructions
Calculator	To use the on-screen calculator, select **Calculator** in the global menu.
Dictionary	To look up definitions and synonyms in the Merriam-Webster dictionary or thesaurus, select **Dictionary** in the global menu.
Formula	To view the on-screen formula sheet, select **Formula** in the global menu.
Help ?	To view the on-screen ***Help Guide*** window, select the question mark ? button in the top-right corner.
Language Toggle	For items and stimuli that have content in two languages, you can select the language in which to display the content for easier readability. By default, items are presented in the language specified at the beginning of the test. To switch the language, select **Language Toggle** in the global menu. For more details, see How to Use the Language Toggle Tool.
Line Reader	To emphasize an individual line of text in a passage or question, select **Line Reader** in the global menu. This tool is not available while the Highlighter tool is in use.
Masking	The Masking tool temporarily covers a distracting area of the test page. To use this tool: • Select **Masking** in the global menu. • Click and drag across the distracting area. • To close the Masking tool, select **Masking** again. To remove a masked area, select in the top-right corner of that area. For more details, see How to Use the Masking Tool.

Tool Name	Instructions
Notes	To enter notes in an on-screen notepad, select **Notes** in the global menu. These notes are available globally and can be accessed from any page in the test. Depending on the tool's settings, advanced editing features, such as word and character count, spell check, and text formatting options may be available. Table 9 lists the different editing tools that may be available.
Pause	To pause a test, select . If you pause the test, then you will be logged out.
Periodic Table	To view the on-screen periodic table, select **Periodic Table** in the global menu.
Print Page	To print the entire test page, select **Print Page** in the global menu.
Print Passage	To print a reading passage, select **Print Passage** in the global menu.
System Settings	To adjust audio volume during the test, select in the top-right corner. Students testing with TTS can also use this tool to adjust TTS settings. Students testing on mobile devices cannot use this tool to adjust volume. To adjust audio volume on mobile devices, students must use the device's built-in volume control. You can also specify the language in which the names of buttons, messages, and other labels are displayed on the test page.
Transcripts	To view a transcript of the audio content for the current test page, select the **Transcripts** button in the global menu. The text in transcripts can be read by screen readers.
Zoom buttons	To enlarge the text and images on a test page, select **Zoom In**. Multiple zoom levels are available. To undo zooming, select **Zoom Out**.

How to Use the Language Toggle Tool (Science Tests Only)

The Language Toggle tool allows students to switch between the two languages in which test content is available. When switching languages, the test content and all the elements on the screen, such as button labels and messages, switch to the selected language. Students can view the content in the chosen language and even respond to items in that language, if applicable.

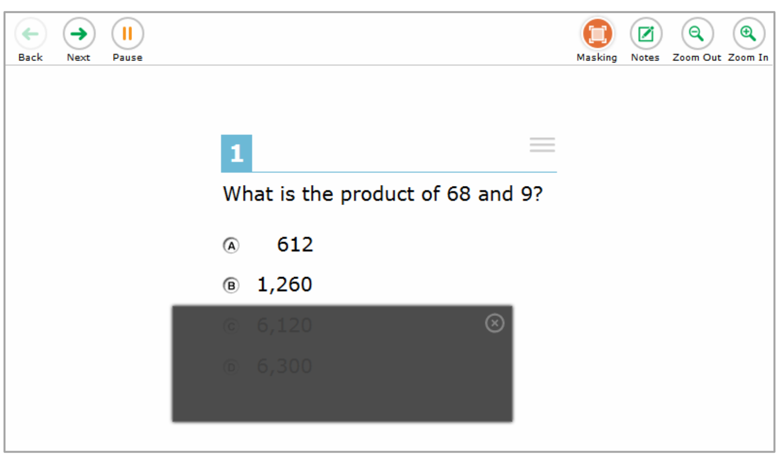
- In order to use the Language Toggle tool, item content must be available in two languages. The default presentation must be set to Spanish and the tool must be enabled in the test settings.
- By default, items are presented in the default language specified in the test settings. If a student toggles the language for an item and then proceeds to the next item, the next item displays in the default language.
- If a student highlights or masks text in one language and then switches to a different language, the corresponding sections will not be highlighted or masked.

How to Use the Masking Tool

The Masking tool allows students to hide distracting areas of the test page.

- To mask an area of a test page:
  a. Select **Masking** in the global menu. The button becomes orange.
  b. Click and drag across the distracting area of the test page. The selected area becomes dark gray. The tool remains active until you deactivate it.
- To deactivate the masking tool, select **Masking** in the global menu again. The button becomes green. Please note that masked areas will remain on the screen until you remove them.

Figure 31. Test Page with Masked Area

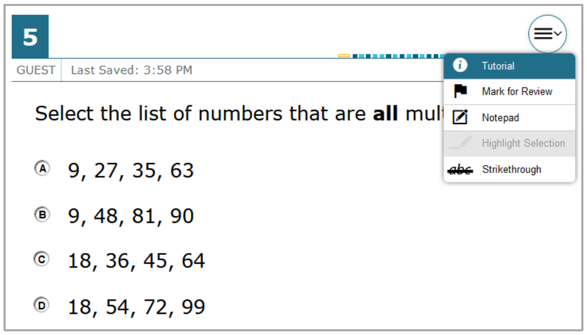


- To remove a masked area from a test page, select in the top-right corner of a masked area.

How to Use Context Menu Tools

A test page may include several elements, such as the question, answer options, and stimulus. The context menu for each element contains tools that are applicable to that element (see Figure 32). Table 4 lists the available context menu tools.

Figure 32. Context Menu for Questions



If a question has multiple response areas, a context menu may be available for each one.

To use the context menu, do one of the following:

- To open the context menu for an item or stimulus, click the context menu or right-click the required elements.
- To open the context menu for an answer option, do one of the following:
  - If you are using a **two-button mouse**, right-click an answer option.
  - If you are using a **single-button mouse**, click an answer option while pressing **Ctrl**.
  - If you are using a **Chromebook**, click an answer option while pressing **Alt**.
  - If you are using a **tablet**, tap the answer option and then tap the context menu button.

Table 4. Context Menu Tools

Tool Name	Instructions
Sign Language	You can watch videos that translate test content into American Sign Language (ASL). • From the context menu, select **American Sign Language**. ▪ If only one ASL video is available, the video opens automatically. ▪ If multiple ASL videos are available, sign language ( ) icons appear next to the test content for each video. Select the icon for the test content you wish to translate into ASL. ▪ If the ASL video content includes chapter markers, you can jump from one point of the video to another.

Tool Name	Instructions
Glossary (English and Translated)	To open the glossary, select a word or phrase that has a border around it. If the TTS feature is enabled, you can select the TTS icon in the Glossary or Word List pop-up window to listen to the content. While the text is being read aloud, the TTS icon is replaced by the speaking icon.
Highlighter	To highlight text, select the text on the screen and then select **Highlight Selection** from the context menu. If multiple color options are available, select an option from the list of colors that appears. To remove highlighting, select **Reset Highlighting** from the context menu. Text in images cannot be highlighted. This tool is not available while the Line Reader tool is in use.
Mark for Review	To mark a question for review, select **Mark for Review** from the context menu. The question number displays a flap in the top-right corner and a flag icon appears next to the question number on the test page. The **Items** pop-up window also displays a flag icon next to the question number.
Notepad	To enter notes for a question, select **Notepad** from the context menu. After entering a note, a pencil icon appears next to the question number on the test page. You can only access your notes for a question on that question's test page. Depending on the tool's settings, advanced editing features, such as word and character count, spell check, and text formatting may be available. Table 9 lists the different editing tools that may be available.
Print Item	To send a print request for an individual question, select **Print Item** from the context menu. After sending the request, a printer icon appears next to the question number on the test page.
Select Previous Version	To view and restore responses previously entered for a Text Response question, select the **Select Previous Version** option from the context menu. A list of saved responses appears. Select the appropriate response and click **Select**. For more information, see How to Use the Select Previous Version Tool.
Strikethrough	For selected-response questions, you can cross out an answer option to focus on the options you think might be correct. There are two options for using this tool: • Option A: a. To activate Strikethrough mode, open the context menu and select **Strikethrough**. b. Select each answer option you wish to strike out. c. To deactivate Strikethrough mode, press **Esc** or click outside the question's response area. • Option B: Right-click an answer option and select **Strikethrough**.

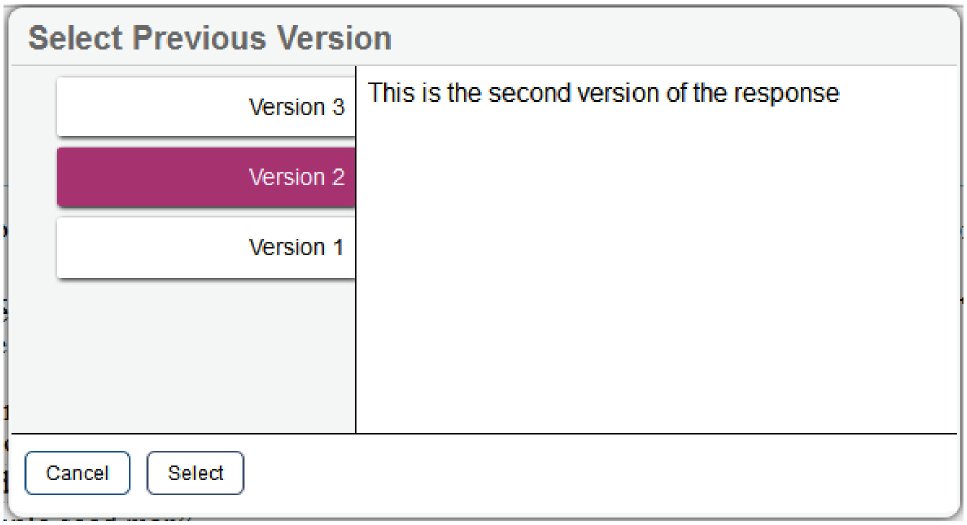
Tool Name	Instructions
Text-to-Speech	To listen to passages and questions, select a **Speak** option from the context menu. To listen to your responses to constructed-response items, use the TTS button in the item response area. For more information, see How to Use the Text-To-Speech Tool.
Tutorial	To view a short video demonstrating how to respond to a particular question type, select **Tutorial** from the context menu.

How to Use the Select Previous Version Tool

The Select Previous Version tool allows students to view and restore responses they previously entered for a text response question. For example, if students type a response, select **Save**, delete the text, and enter new text, they can use this tool to recover the original response. Please note that if the student's test pauses, any responses entered prior to pausing cannot be recovered.

1. To recover a previously entered response, select the **Select Previous Version** option from the context menu. The ***Select Previous Version*** window appears, listing all the saved responses for the question in the left panel.
2. Select a response version from the left panel. The text associated with that response appears in the right panel.
3. Select the **Select** button. The selected response appears in the text box for the question.

Figure 33. Select Previous Version Window



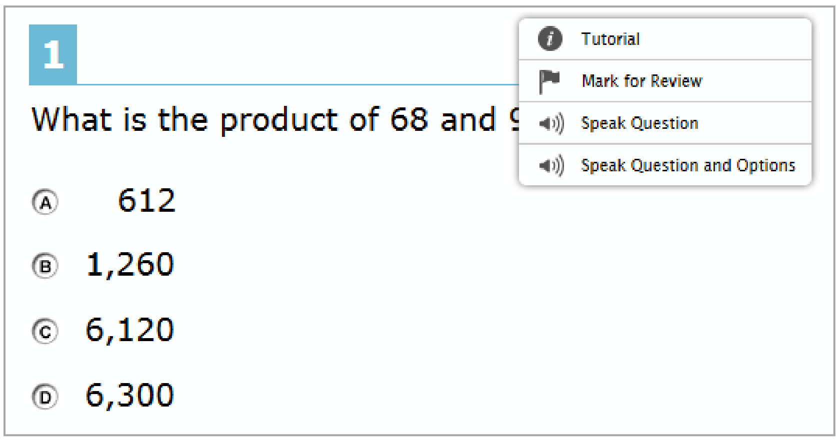
How to Use the Text-To-Speech Tool

Students testing with TTS can listen to passages, questions, and answer options using the TTS options available in the selected element's context menu. Students may also listen to their responses to constructed-response items based on their accommodations. A test may include text that is not allowed to be read aloud. This text will be skipped when using TTS.

TTS is available when using the Secure Browser or a supported Chrome, Firefox, or Edge browser.

If a student is using Text-to-Speech, tracking is enabled so that the words become highlighted as they are read aloud.

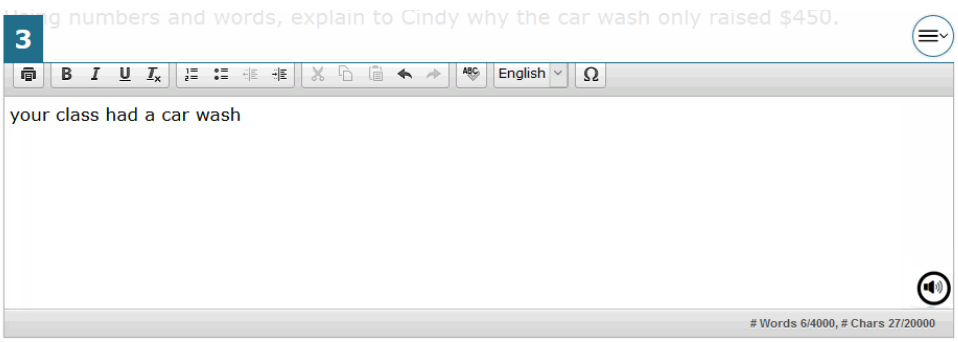
Figure 34. TTS Options for Questions



- To listen to a passage, students open the passage context menu and select a **Speak** option. Students can also select a portion of text to listen to, such as a word or phrase. To do this, students select the text, open the passage context menu, and select **Speak Selection**.
  - When listening to passages, students can pause TTS and then resume it at the point where it was paused. While this functionality is not available on ChromeOS. Students testing on a ChromeOS device can resume a paused TTS passage by selecting the remaining text to be read aloud and selecting **Speak Selection** from the context menu.
- To listen to a question or answer options, students open the question context menu and select one of the following **Speak** options:
  - To listen only to the question, students select **Speak Question**.
  - To listen to a multiple-choice question and all answer options, students select **Speak Question and Options**.
  - To listen only to an answer option, select **Speak Option** from the context menu and then select the answer option. Students can also right-click the answer option and select **Speak Option**.

- Some tests may also display ear icons by questions and answer options when TTS is enabled. Students can click these icons to listen to the corresponding test content.
- Depending on their accommodations, students may also be able to use TTS to listen to their responses to constructed-response items. Students can select or as available in the formatting toolbar of the text-response area to listen to the text as entered.

Figure 35. Constructed-Response Item with TTS Tool



How to Use Other Tools

In addition to the global tools and context menu tools, there are some additional tools that may be available to students based on their accommodations or the test page layout. Table 5 lists the additional tools available in the Student Testing Site.

Table 5. Other Tools

Tool Name	Instructions
Closed Captioning	Questions and stimuli with audio elements automatically display closed captions for students testing with the appropriate accommodations. You can select the up arrow to move the closed captioning to the top of the screen or the down arrow to move it to the bottom. You can also exit closed captioning by selecting .
Expand Passage	To expand the passage section, select the double arrow icon. The stimulus section will expand and overlap the question section for easier readability. To collapse the expanded section, select the double arrow icon again.
Expand buttons	You can expand the stimulus section or the question section for easier readability. • To expand the stimulus section, select the right arrow icon below the global menu. To collapse the expanded stimulus section, select the left arrow icon in the top-right corner. To expand the question section, select the left arrow icon below the global menu. To collapse the expanded question section, select the right arrow icon in the top-left corner.

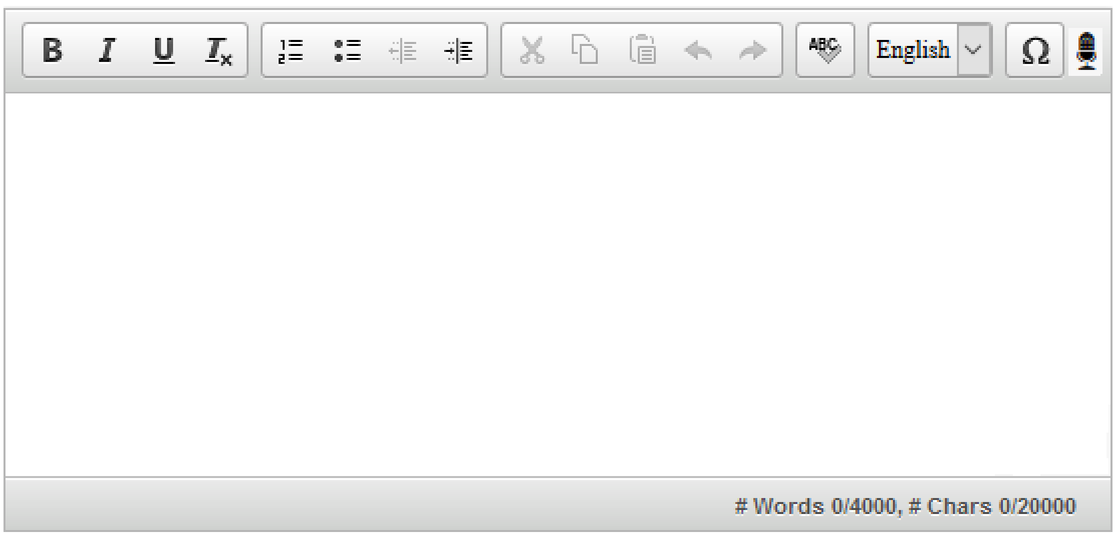
Tool Name	Instructions
Speech-to-Text (STT)	Speech-to-Text allows students testing with the appropriate accommodations to dictate responses to constructed-response items. To use STT, select the microphone icon in the formatting toolbar of the item response area and begin speaking. The dictated response will be transcribed in the item response area. For more information, see the section How to Use the Speech-to-Text Tool.

How to Use the Speech-to-Text Tool

Students with the appropriate accommodations may use the Speech-to-Text (STT) tool when responding to text response items. In supported items, the STT tool allows a student to dictate a spoken response that is transcribed in the item response area.

To begin dictating, the student selects the microphone button in the formatting toolbar of the item's text response area. Students with the appropriate accommodations can also select whether to dictate text in English or Spanish from the language drop-down list in the formatting toolbar.

Figure 36. Text-Response Area with STT Tool



As the student speaks, the words are transcribed in the text response area. There may be a slight delay while the text is being transcribed. Dots appear in the text response area to indicate that the transcription is in process.

The student can stop the dictation by selecting the stop button that appears while dictating. Note that the button automatically reverts to the microphone button if no sound is detected for a specified period. Students can click the microphone button again to resume dictation. Students can dictate for five minutes at a time.

Depending on the tool settings, the entered text may be auto-punctuated. Students can also control the punctuation and grammar of the text through speech commands to some extent. For example, students can say, "New Paragraph" to create a new paragraph. It is ultimately the student's responsibility to ensure the accuracy of the transcription, as well as grammar and punctuation.

The buttons in the formatting toolbar are disabled while dictation is on. The buttons reactivate once students stop the recording. Students cannot navigate away from the test page while dictation is on.

How Students Complete a Test

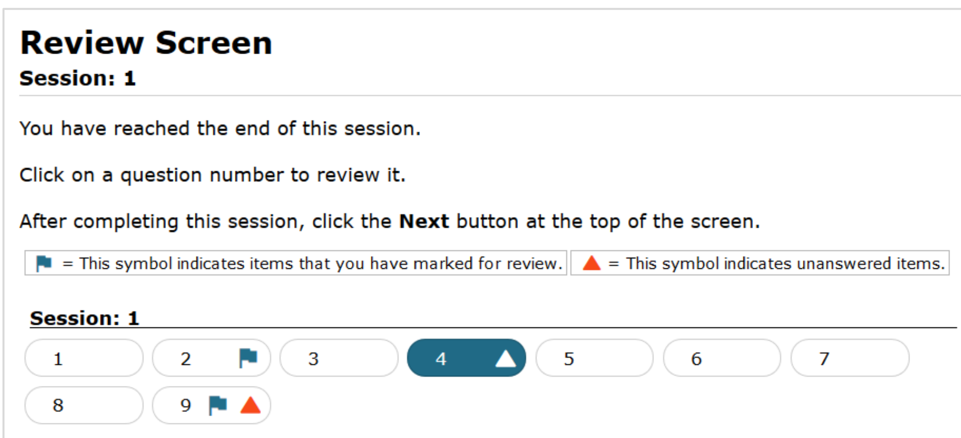
After students complete their test, they need to submit their test for scoring.

How to Complete a Test Segment

In segmented tests, the ***End Segment*** page appears after students finish the last question in a segment. This page allows students to review questions from the current segment (and earlier segments, if allowed) or proceed to the next segment.

- To review questions, students select a question number.
  - A flag icon appears for any questions marked for review. A warning icon appears for any unanswered questions.
- To move to the next segment, students select **Next** in the global menu.

Figure 37. End Segment Page



Please note that if the test blocks access to completed segments, students cannot return to the segment after selecting **Next**.

How to Submit a Test

To complete the testing process, students must submit their tests when they are finished answering questions.

Once students submit their tests, they cannot return to the test or modify answers.

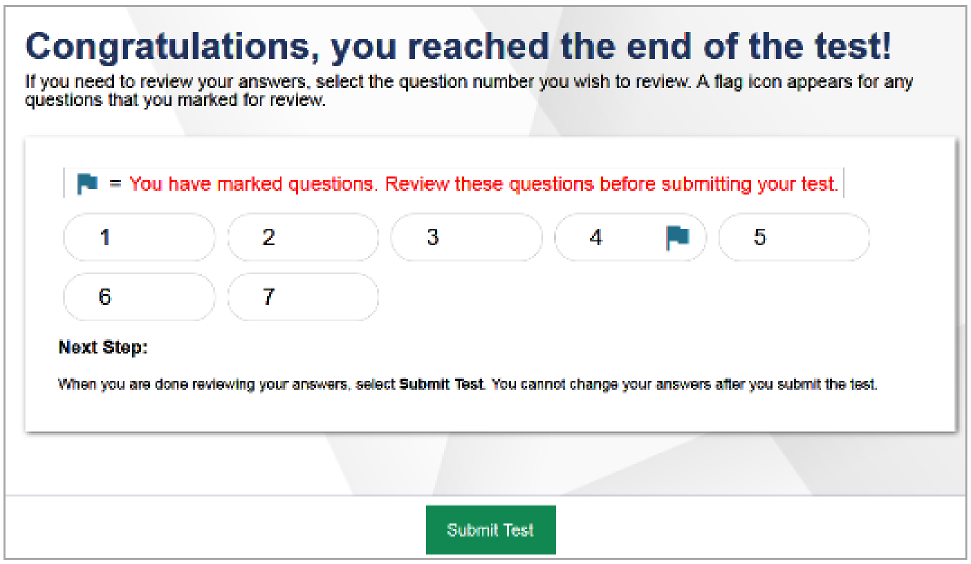
1. To end a test, students select **End Test** in the top-left corner, which appears after students respond to the last test question. A confirmation message appears.

Figure 38. Global Menu with End Test Button

2. Students select **Yes**. The ***End Test*** page appears, allowing students to review answers and submit the test for scoring.

   – A flag icon appears for any questions marked for review. A warning icon appears for any unanswered questions (students can only skip items on the Alternate Montana Science Assessment).

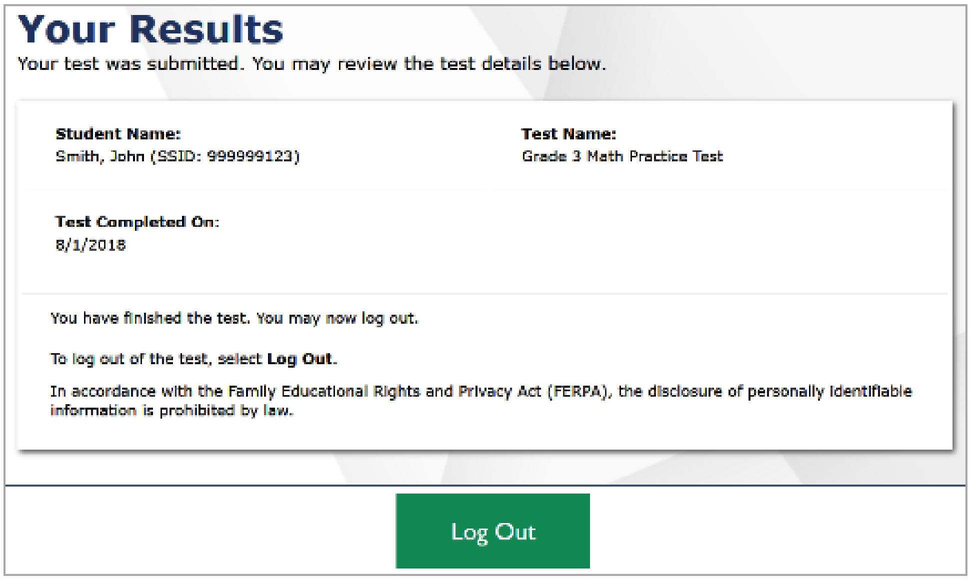
Figure 39. End Test Page



3. *Optional*: To review previous answers, students select a question number. When finished reviewing, they can return to the ***End Test*** page by selecting **End Test** again.

4. To submit the test, students select **Submit Test**, then select **Yes** in the confirmation message that appears. The ***Your Results*** page appears, displaying the student's name, the test name, and the completion date.

Figure 40. Your Results Page



5. To exit the Student Testing Site, students select **Log Out,** and then close the Secure Browser.

Appendix

A

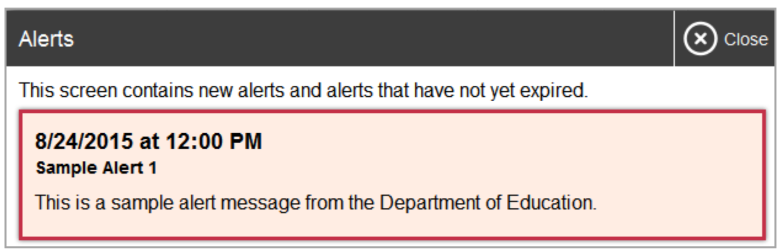
Alert Messages

The Montana Office of Public Instruction can send statewide alerts that appear as pop-up messages on the TA Site.

Figure 41. Alerts Button

- To view alerts, select **Alerts** in the banner. The ***Alerts*** window appears, listing all the active alert messages.

Figure 42. Record of Alerts



Alternate Assessments in the Student Testing Site

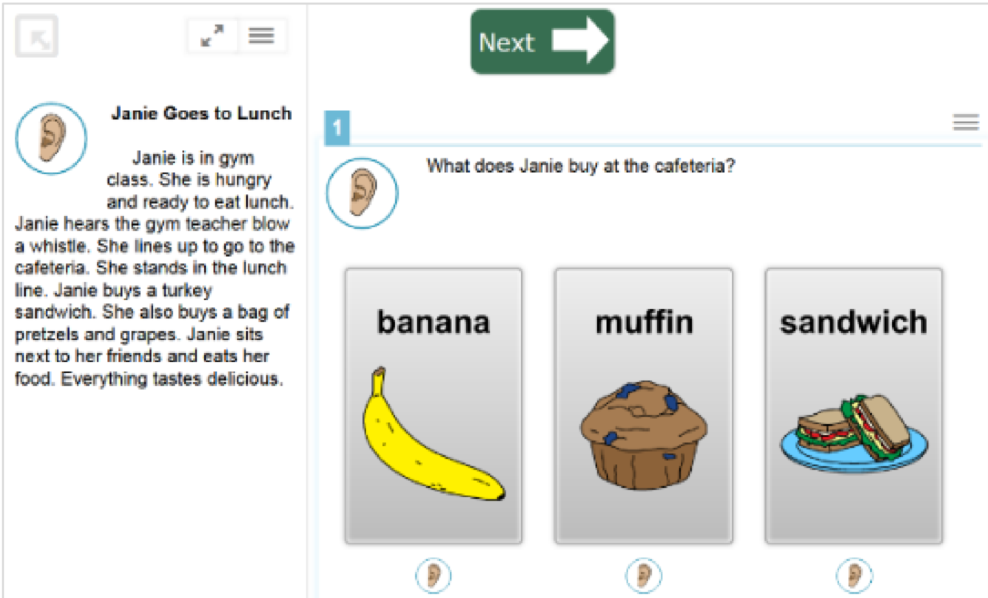
You can administer alternate assessments online by adding the tests to sessions in the TA Site following the same procedure used to add any other test to a session. However, the appearance and behavior of the Student Testing Site is slightly different for alternate assessments.

Navigating Alternate Assessments

The Student Testing Site displays alternate assessments in full-screen mode with the global menu hidden by default.

- To proceed to the next question, students select Next.
- To exit full-screen mode and access the global menu, students select in the upper-left corner.

Figure 43. Alternate Assessment Layout



- To return to full-screen mode, students select **Full Screen** in the global menu.

Figure 44. Full Screen Button

Listening to Audio Content

Some test pages display ear icons, which read aloud the content for passages, questions, and answer options. These icons also appear when students take alternate assessments with TTS enabled (in addition to the TTS options available in the context menu).

To listen to the test's content, students must do the following:

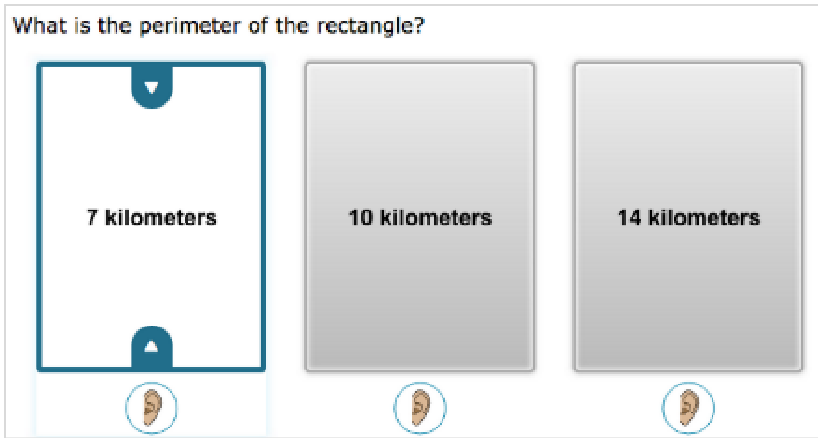
- To listen to a passage, select beside the passage.
- To listen to a question and its answer options, select beside the question.
- To listen only to an answer option, select below the answer option.
- To stop the audio, click while it is playing.
  - You cannot pause audio content. If you select again, it will play the audio from the beginning.

Responding to Questions and Ending a Test

Alternate assessment questions display answer options as cards, which may include text, graphics, or both.

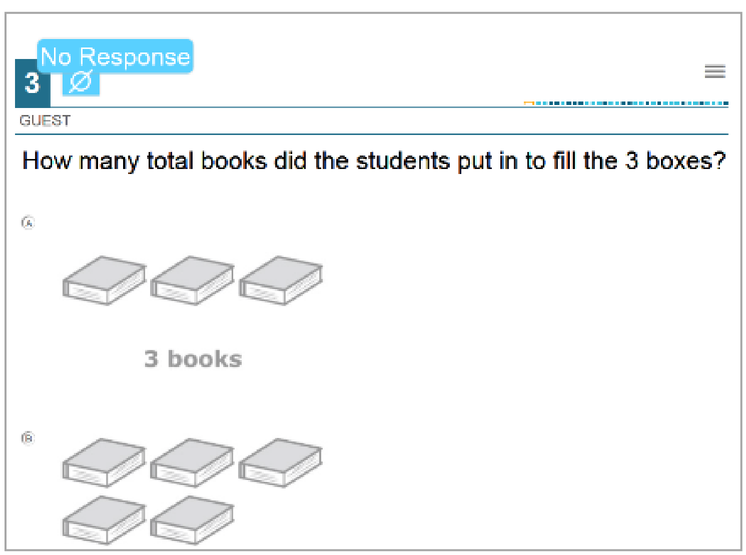
- To select an answer option, students select a card. Once a card is selected, arrows appear at the top and bottom of the card, along with a blue border.

Figure 45. Alternate Assessment Card Item



- To indicate that no response or an unsupported response was provided for an item, do the following:
  - From the item's context menu, select **Mark as No Response** or **Mark as Unsupported Response**, then select **Yes** in the confirmation message that appears.
    - The item's response options disable and a null symbol ∅ displays next to the item number, in the **Items** drop-down list, and on the ***Review*** page at the end of the test. For reporting purposes, the item is assigned a condition code of NR for no response or UR for unsupported response.

Figure 46. Item Marked as No Response



   - *Optional*: To revert an item marked as no response or unsupported response, select **Unmark No Response** or **Unmark Unsupported Response** from the context menu. If you previously entered a response, that response will not be retrieved.

- To end a test, exit full-screen mode and select **End Test**. You can submit the alternate assessment by following the same procedure used to submit all other assessments in the Student Testing Site.

E

Expiration Rules for Test Opportunities

Opportunities refer to the number of times a student can take a test within a range of dates. Tests may have one opportunity or multiple opportunities. A student's test opportunity remains active until the student submits the test or until the opportunity expires. Once a test opportunity expires, the student cannot complete or review the test.

K

Keyboard Commands in the Student Testing Site

Students can use keyboard commands to navigate between test elements, features, and tools. Some important things to note about keyboard commands are:

- Keyboard commands require the use of the primary keyboard, so please do not use keys in a numeric keypad.
- Some keyboard commands (such as the commands for using the Line Reader) may not work when testing on iPadOS devices connected to an external keyboard.
- When Permissive Mode is enabled for a test, keyboard commands are blocked and will not work.

Keyboard Commands for Sign-In Pages and In-Test Pop-Ups

Table 6 lists keyboard commands for selecting options on the sign-in pages or pop-up windows that appear during a test.

Table 6. Keyboard Commands for Sign-In Pages and Pop-Up Windows

Function	Keyboard Commands
Move to the next option	**Tab**
Move to the previous option	**Shift + Tab**
Select the active option	**Enter**
Mark checkbox	**Space**
Scroll through drop-down list options	**Arrow Keys**
Close pop-up window	**Esc**

Keyboard Commands for Test Navigation

Table 7 lists keyboard commands for navigating tests and responding to questions.

Table 7. Keyboard Commands for Test Navigation

Function	Keyboard Commands
Scroll up	**Up Arrow**
Scroll down	**Down Arrow**
Scroll to the right	**Right Arrow**
Scroll to the left	**Left Arrow**
Move to the next element	**Tab**
Move to the previous element	**Shift** + **Tab**
Select an answer option	**Space**
Go to the next test page	**Ctrl** + **Right Arrow**
Go to the previous test page	**Ctrl** + **Left Arrow**
Open the global menu	**Ctrl** + **G**
Open a context menu	**Ctrl** + **M**

Keyboard Commands for Global and Context Menus

Students can use keyboard commands to access tools in the global and context menus. For more information about tools in these menus, see the section How Students Use Test Tools.

Global Menu

1. To access the global menu tools using keyboard commands, press **Ctrl** + **G**. The global menu list opens.
2. To move between options in the global menu, use the **Up** or **Down** arrow key.
3. To select an option, press **Enter**.
4. To close the global menu without selecting an option, press **Esc**.

Context Menus

1. To open the context menu for an element (question, answer option, or stimulus), navigate to the element using the **Tab** or **Shift** + **Tab** command.
2. Press **Ctrl** + **M**. The context menu for the selected element opens.
3. To move between options in the context menu, use the **Up** or **Down** arrow keys.
4. To select an option, press **Enter**.

5. To close the context menu without selecting an option, press **Esc**.

Keyboard Commands for Highlighting Selected Regions of Text

This section explains how to use keyboard commands to select a text excerpt (such as a word in a passage) and highlight it. These instructions apply only to students using the Secure Browser.

1. To select text and highlight it, navigate to the element containing the text you want to select.

2. Press **Ctrl** + **M** to open the context menu and navigate to **Enable Text Selection**.

3. Press **Enter**. A flashing cursor appears at the upper-left corner of the active element.

4. To move the cursor to the beginning of the text you want to select, use the arrow keys.

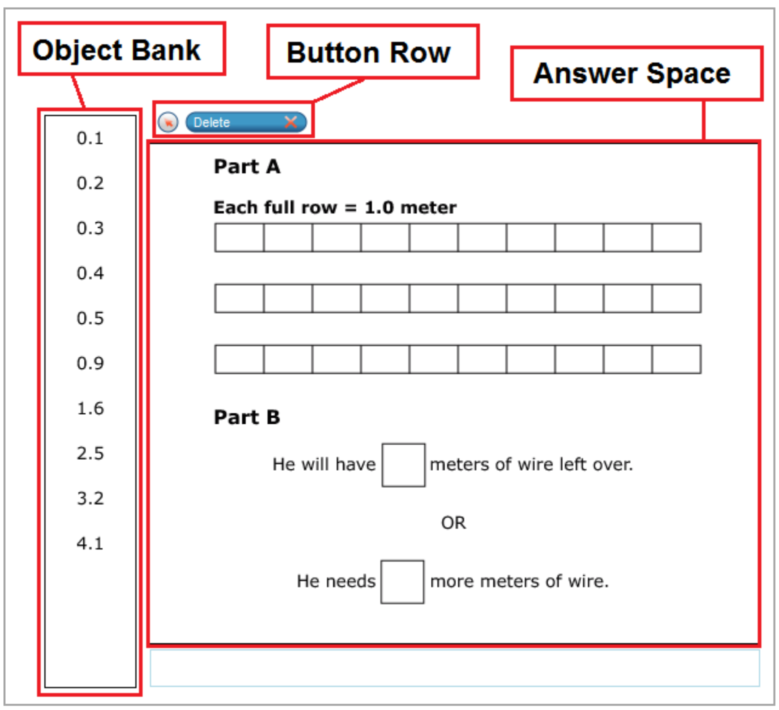
5. Press **Shift** and an arrow key to select your text. The text you select appears shaded.

6. Press **Ctrl** + **M** and select **Highlight Selection**.

Keyboard Commands for Grid Questions

Questions with the grid response area may have up to three main sections: an answer space, which is the grid area where students enter the response; an object bank, which is a panel containing objects you can move to the answer space; and a button row, which appears above the answer space and may include **Delete**, **Add Point**, **Add Arrow**, **Add Line**, **Add Circle**, **Add Dashed Line**, and **Connect Line** buttons.

Figure 47. Grid Question



- To move between the main sections, do the following:
  - To move clockwise, press **Tab**. To move counterclockwise, press **Shift** + **Tab**.
- To add an object to the answer space**,** do the following:
  a. With the object bank active, use the arrow keys to move between objects. The active object has a blue background.
  b. To add the active object to the answer space, press **Space**.
- To use the action buttons, do the following:
  a. With the button row active, use the left and right arrow keys to move between the buttons. The active button is white.
  b. To select a button, press **Enter**, and then press **Space** to apply the point, arrow, or line to the answer space.
- To move objects and graph elements in the answer space, do the following:
  a. With the answer space active, press **Enter** to move between the objects, and then press **Space**. The active object displays a blue border.
  b. Press an arrow key to move the object. To move the object in smaller increments, hold **Shift** while pressing an arrow key.

Keyboard Commands for Equation Questions

Equation questions allow students to use keyboard commands to open a menu listing the special characters they can insert into the response area.

1. To insert special characters in the response area, with the focus in the text field of the response area, press **Alt** + **7**. The ***Special Characters*** window opens.
2. To move between options in the context menu, use the **Up** or **Down** arrow keys.
3. To add the selected option to the response area, press **Enter**.

L

Login Information for the TA Site

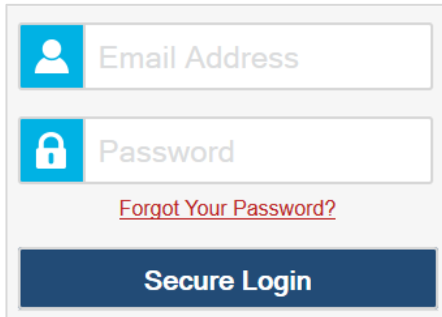
To access the TA Sites, your TIDE administrator must first create your account in TIDE. Once your account is created, you receive an account activation email. You can log in to the TA Sites after activating your account.

1. Navigate to the Montana Testing portal (www.mt.portal.cambiumast.com).
2. Select the appropriate assessment card. Then select the appropriate TA Site:
   a. To access the TA Interface, select **Test Administration**.

b. To access the TA Training Site, select **Training Test Administration**.

3. The ***Login*** page appears. Enter your email address and password.

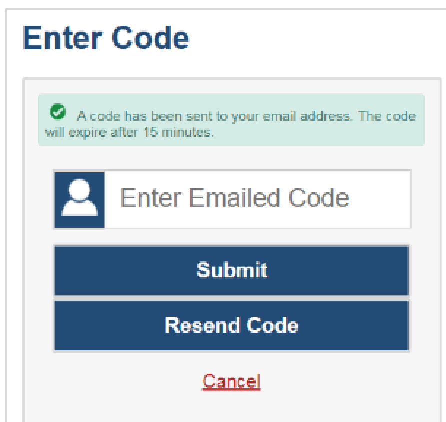
Figure 48. Login Page



4. Select **Secure Login**. The selected TA Site appears.

a. If the ***Enter Code*** page appears, an authentication code will be emailed to you. You must enter this code in the *Enter Emailed Code* field and click **Submit** within 15 minutes. If the code expires, click **Resend Code** to request a new one.

Figure 49. Enter Code Page



b. If you receive a warning message about the Test Security Agreement (TSA), you must sign the TSA in TIDE and then log in again.

c. If you receive a warning message about not being certified, you must complete the TA Certification Course and then log in again.

5. If you are associated with multiple institutions that have test windows set, a message prompts you to select a testing institution. Select your institution and select **Go**. To change the institution, you must log out and then log back in.

P

Pause and Test Timeout Rules

Pause Rules

TAs and students can pause a test in order to log out the student from the test session temporarily. Students cannot review or modify answered questions after their test pauses for more than 20 minutes, even if they marked questions for review. The only exceptions to this rule are if a student pauses the test before answering all the questions on the current page or if you submit an appeal in TIDE.

These pause rules apply regardless of whether the test is paused by the student or TA, or if it was paused due to a technical issue.

Test Timeout Rules

A warning message displays after 20 minutes of test inactivity. Students who do not click **OK** within 30 seconds after this message appears are logged out. This timeout automatically pauses the test.

Practice Test Site Student Sign-in Process

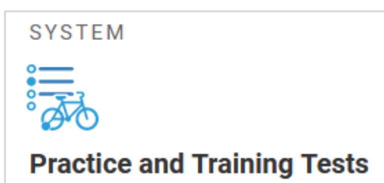
The Student Training Site allows students to take practice tests. Aside from the sign-in process, the Practice Test Site has the same appearance and functionality as the Student Testing Site. For information about how to sign in to the Student Testing Site, see How Students Sign in and Select Tests.

Students can take practice tests in proctored sessions created in the TA Training Site or in non-proctored (guest) sessions. Students also have the option to sign in to the test sessions with their real identities to take tests specific to their grades or sign in as guests to take tests for any grade level.

1. To access the Student Training Site, do one of the following:

   – From the Montana Testing portal (www.mt.portal.cambiumast.com), select the **Practice and Training Tests** card.

Figure 50. Practice and Training Tests Card



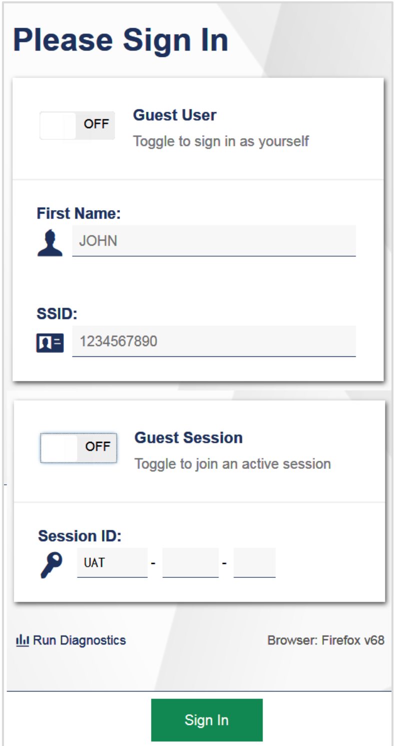
   – Launch the Secure Browser and select the **Take a Practice Test** button.

2. To sign in, students do the following:

   – To sign in as a guest, students set the **Guest User** toggle to **On**. Otherwise, to use their real credentials, students set the **Guest User** toggle to **Off** and then enter their first name and SSID.

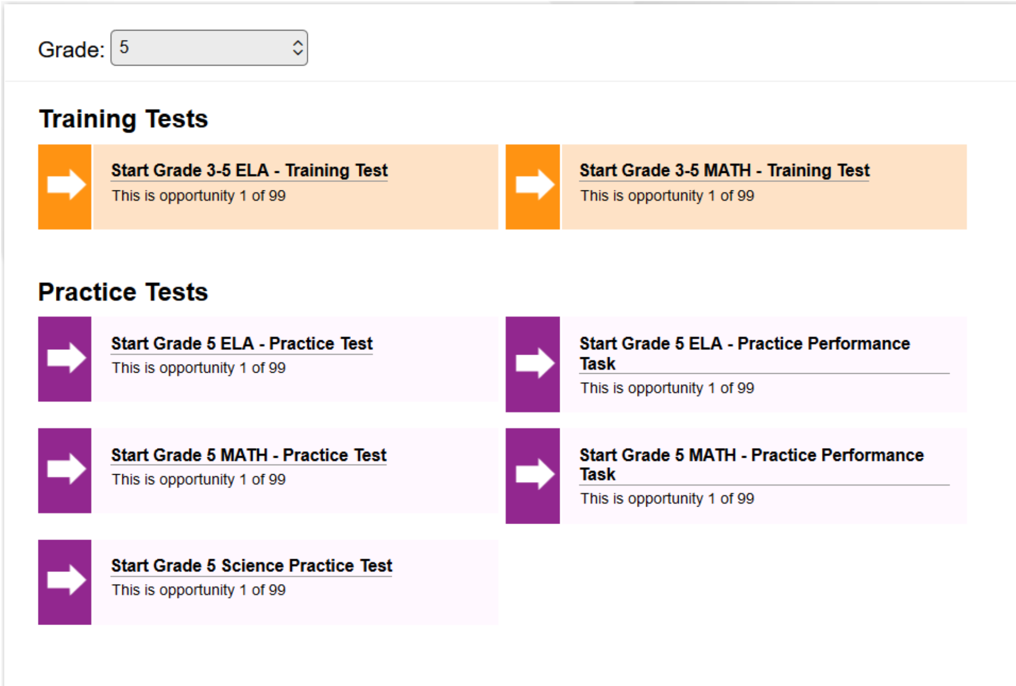
- To join a guest session, students set the **Guest Session** toggle to **On**. Or else, to join a proctored session, students set the **Guest Session** toggle to **Off** and enter the Session ID from the TA Training Site.
- Students select **Sign In**.
  - If signed in with their real identities, the ***Is This You*** page appears. Students verify their information and click **Yes** to proceed to the ***Your Tests*** page.
  - If signed in as guest users, students are taken directly to the ***Your Tests*** page.

Figure 51. Student Training Site Login Page



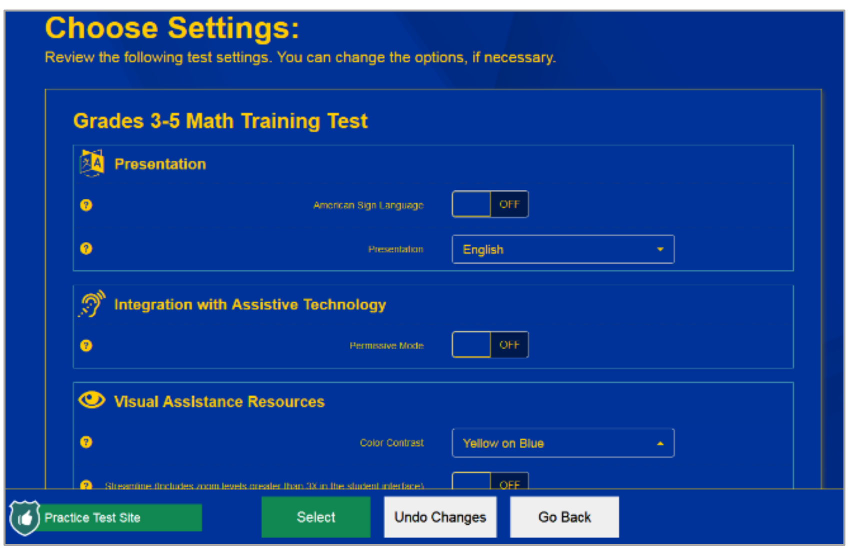
3. On the ***Your Tests*** page, students do one of the following:
   - If signed in with their real identities, students select a test from the ones available for their grade.
   - Students signed in as guests select their grade level from the drop-down list to view the tests available for that grade and then select a test.

Figure 52. Your Tests Page



4. If the students signed in to a guest session, they must select the test settings they wish to use from the ***Choose Settings*** page and then select the **Select** button. When selecting the color of the text and background, mouse-pointer, and print size settings, students can see a live preview of their selected settings.

Figure 53. Choose Settings Page



5. If the test includes audio content or TTS settings, the ***Audio/Video Checks*** page appears, displaying the functionality checks that need to be performed. Students must follow the instructions on this page to ensure their device is working properly.

6. On the final sign-in page, students may review the help guide, their test settings, and the additional test information, then select **Begin Test Now** to start or resume their test opportunity.

Print Session Information

You can print a snapshot of the TA Site as it currently appears if you wish to keep a hard-copy record of the Session ID or list of approved students. Please note that federal law prohibits the release of students' personally identifiable information. All printouts must be securely stored and then destroyed when no longer needed.

1. In the TA Site banner, select **Print Session**. The computer's print dialog window appears.

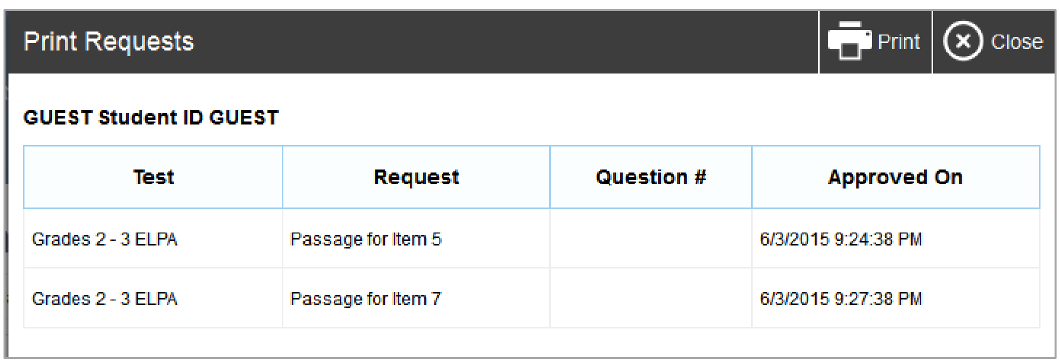
2. Select **OK**.

Print Approved Requests Information

You can view and print a list of every print request you approved for students during the current session. For information about how to approve a print request, see How to Approve a Student's Print Request.

1. In the TA Site banner, select **Approved Requests**. The ***Print Requests*** window appears, listing all the approved print requests.

Figure 54. Print Requests Window



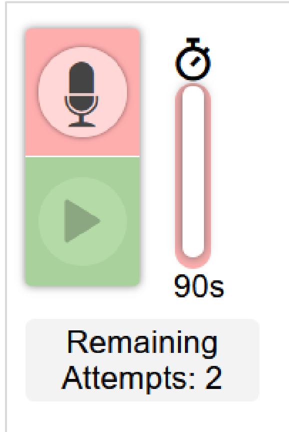
2. Select **Print** to print the list.

R

Recording Spoken Responses

Some test items may require students to record a spoken response. In order to record a spoken response, students must test with a headset that has a functioning microphone.

Figure 55. Spoken Response Item



If the Student Testing Site detects that there may be audio issues with a student's recorded response, a warning message appears, encouraging them to play the response they entered to make sure it is audible.

If a spoken response item limits the number of attempts students can make to record a response, the Remaining Attempts section indicates how many attempts the student has left. Recording a new attempt erases the previous attempt. If a student runs out of attempts, the last attempt they recorded will be used.

If a spoken response item limits the length of a student's recording, a timer appears next to the microphone icon. The timer begins counting down when students start recording their response. The recording automatically stops when the timer runs out.

S

Secure Browser

The Secure Browser ensures test security by prohibiting access to external applications and navigation away from the test. When the Secure Browser launches, it checks for other applications running on the device. If it detects a denylisted application, it displays a message listing the offending application and prevents the student from testing. This also occurs if a denylisted application launches while the student is already in a test.

In most cases, a detected denylisted application is a scheduled or background job, such as anti-virus scans or software updates. The best way to prevent denylisted applications from running during a test is to schedule such jobs outside of planned testing hours.

There are additional measures you can implement to ensure the test environment is secure:

- **Close External User Applications**

  Before launching the Secure Browser, or prior to administering the online tests, close all non-required applications on testing devices, such as word processors and web browsers.

- **Avoid Testing with Dual Monitors**

  Students should not take online tests on computers connected to more than one monitor. Systems that use a dual-monitor setup typically display an application on one screen while another application is accessible on the other screen.

- **Disable Screen Savers and Timeout Features**

  On all testing devices, be sure to disable any features that display a screensaver or log users out after a period of inactivity. If such features activate while a student is testing, the Secure Browser logs out the student from the test.

Using the Secure Browser with Accessibility Software

Permissive Mode is an accommodation that allows students to use non-embedded assistive technology to complete tests in the Secure Browser. It must be turned on for any students testing with third-party assistive technology tools. When Permissive Mode is turned on, the Secure Browser's security settings will be partially lowered to allow students to use tools that would otherwise be blocked. This accommodation should be assigned to students in TIDE before they begin testing.

Permissive Mode is available for computers running supported desktop Windows and macOS operating systems only. Permissive Mode on macOS devices is supported only with Secure Browser version 12.5. When using Windows devices, the task bar remains on-screen throughout the test after enabling accessibility software. However, forbidden applications are still prohibited.

When Permissive Mode is turned on, standard keyboard commands in the Secure Browser will be disabled in order to accommodate any potential keyboard commands associated with the assistive technology the student may be using.

Please note that accessibility software must be certified for use with the Online Testing System and forbidden applications will still not be allowed to run. For information about supported software and operating systems, see the *Assistive Technology Manual*.

Permissive Mode activates when the student is approved for testing. Students who have the Permissive Mode setting enabled should not continue with the sign-in process until their accessibility software is correctly configured.

1. To use accessibility software with the Secure Browser, open the required accessibility software.

2. Open the Secure Browser. Begin the normal sign-in process up to the TA approval step.

3. When a student is approved for testing, the Secure Browser allows the operating system's menu and task bar to appear.

   – **Windows**: On Windows, the Secure Browser resizes, and the taskbar remains visible inside the test in its usual position. Students can press **Alt** + **Tab** to switch between the Secure Browser and accessibility applications, such as JAWS and NVDA, that they are permitted to use in their test session. The task bar remains on-screen after enabling accessibility software. However, denylisted applications are still prohibited.

- **macOS**: On macOS, the Secure Browser resizes, and students can view the dock in its usual position inside the test. If the dock is set to autohide, no resizing occurs, and the dock is only visible when the mouse is moved toward the bottom of screen. Students can press **Cmd** + **Tab** to switch between the Secure Browser and permitted accessibility applications.

4. The student must immediately switch to the accessibility software that is already open on the computer so that it appears over the Secure Browser. The student cannot click within the Secure Browser until the accessibility software is configured.

- **Windows:** To switch to the accessibility software application, click the application in the task bar.

- **macOS:** To switch to the accessibility software application, click the application in the dock.

5. The student configures the accessibility software settings as needed.

6. After configuring the accessibility software settings, the student returns to the Secure Browser and continues with the sign-in process. At this point, the student can no longer switch back to the accessibility software. If changes need to be made, the student must sign out and then sign in again.

If Permissive Mode is turned off, the Secure Browser reoccupies the whole screen so that the taskbar or dock is no longer visible, and the student's ability to switch between any applications and the Secure Browser is suppressed.

Accessing the Secure Browser on Mobile Devices

Tablets and Chromebooks should be configured for testing before you provide them to students. For more information, see the OS and platform specific technical documents on the State Assessment Program Portal.

To configure iPadOS devices:

- Tap the **SecureTestBrowser** icon.

To configure Chromebooks:

- From the **Apps** link on the ChromeOS login screen, select **SecureTestBrowser**.

Closing the Student Testing Site on Tablets

After a test session ends, close the **SecureTestBrowser** application on student tablets.

To close the Student Testing Site on iPadOS devices:

1. Double-tap the Home button. The multitasking bar appears.

2. Locate the **SecureTestBrowser** app preview and slide it upward.

To close the Student Testing Site on Chromebooks:

- Click **Close Secure Browser** in the top-right corner.

Force-Quitting the Secure Browser

In the rare event that the Secure Browser becomes unresponsive, you can force-quit the Secure Browser. Please note that the Secure Browser hides features such as the macOS dock. If the Secure Browser is not closed correctly, then the dock may not reappear correctly, requiring you to reboot the device. Avoid using a force-quit command if possible.

To force the Secure Browser to close, use the keyboard command for your operating system as shown in Table 8. This action logs out the student from the test. When the Secure Browser is opened again, the student logs back in to resume testing.

Table 8. Force Quit Secure Browser Keyboard Commands

Operating System	Key Combination
Windows*	• **Ctrl** + **Alt** + **Shift** + **F10**
macOS*	• **Ctrl** + **Alt** + **Shift** + **F10**. The **Ctrl** key may appear as **Control**, **Ctrl**, or **^**
Linux	• **Ctrl** + **Alt** + **Shift** + **Esc**

* If you are using an Apple keyboard, you may need to press **Ctrl** + **Shift** + **Option** + **F10.** If you are using a laptop or notebook, you may also need to press **Function** before pressing **F10**.

Force-quit commands do not exist for the Secure Browser for iPadOS, and ChromeOS devices.

- **iPadOS:** Double-tap the Home button, then close the app as you would any other iPadOS app.
- **ChromeOS:** To exit the Secure Browser from the sign-in screens, press **Ctrl** + **Shift** + **S**. You cannot force-quit once the test begins.

Student Lookup Feature

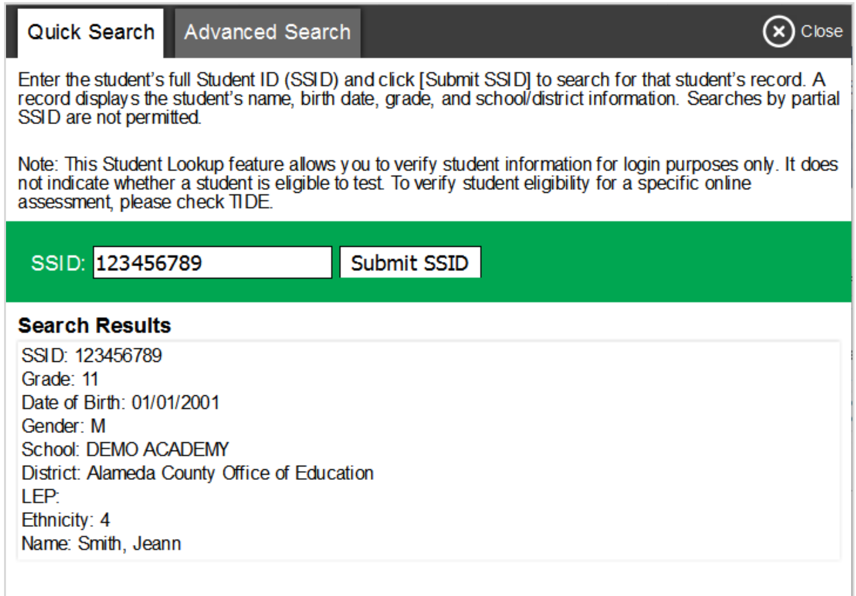
You can use the student lookup feature in the TA Site to perform a quick or advanced search for student information. This is useful if students signing in to your test session cannot remember their login information.

1. To perform a quick search:

    a. In the banner, select **Student Lookup**.

    b. Enter a student’s full SSID and click **Submit SSID**. Search results appear below the search field.

Figure 56. Student Lookup: Quick Search



2. To perform an advanced search:

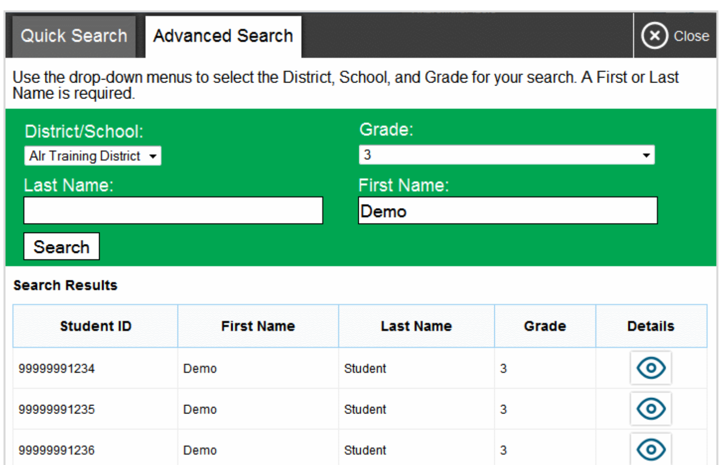
a. In the banner, select ***Student Lookup,*** and then select **Advanced Search**.

b. Fill out the search fields with the necessary information including district/school, first and last name (exact matches only), date of birth, and enrolled grade.

c. Select ***Search***. Search results appear below the search fields.

d. To view a student's information, click in the Details column.

Figure 57. Student Lookup: Advanced Search



T

Text-Response Questions

Text-response item types in the Student Testing Site may allow students to use a formatting toolbar. This toolbar is available above the response field (see Figure 58) and also appears when students right-click in the text area. The formatting toolbar allows students to apply styling to text and use standard word-processing features. Table 9 provides an overview of the formatting tools available.

Figure 58. Text Response Question with Formatting Toolbar

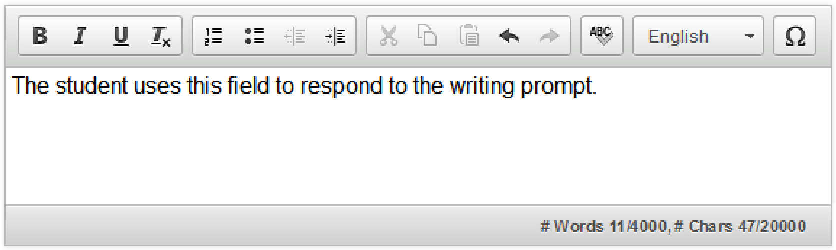
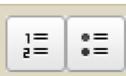
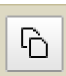
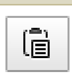
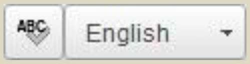


Table 9. Description of Formatting Tools

Tool	Description of Function
	Print the entered text.
	Enable speech-to-text (dictation).
	Bold, italicize, or underline selected text.
	Remove formatting that was applied to the selected text.
	Insert a numbered or bulleted list.
	Indent a line of selected text.
	Decrease indent of text.
	Cut selected text.
	Copy selected text.
	Paste copied or cut text.
	Undo the last edit to text or formatting in the response field.
	Redo the last undo action.

Tool	Description of Function
	Use spell check to identify potentially misspelled words in the response field. The drop-down list allows you to set a language for this tool.
	Add special characters in the response field.

Spell Check Feature

The spell check tool identifies words in the response field that may be misspelled.

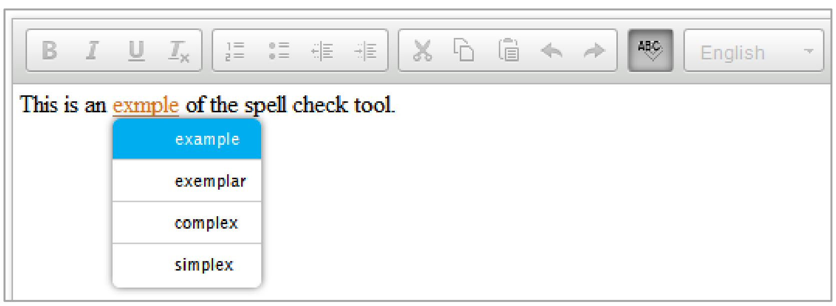
1. Select a language for the spell check tool from the Spell Check drop-down list, if necessary.

Figure 59. Spell Check Drop-Down List

2. In the toolbar, select . Potentially incorrect words change color and become underlined.

3. Select a misspelled word. A list of suggestions appears.

Figure 60. Spell Check Tool



4. Select a replacement word from the list. If none of the replacement words are correct, close the list by clicking anywhere outside it.

5. To exit spell check, select  again.

Special Characters Feature

Students can add mathematical, accented, and other symbols.

1. To add a special character, in the toolbar, select .

2. In the window that pops up, select the required character.

Transfer a Test Session

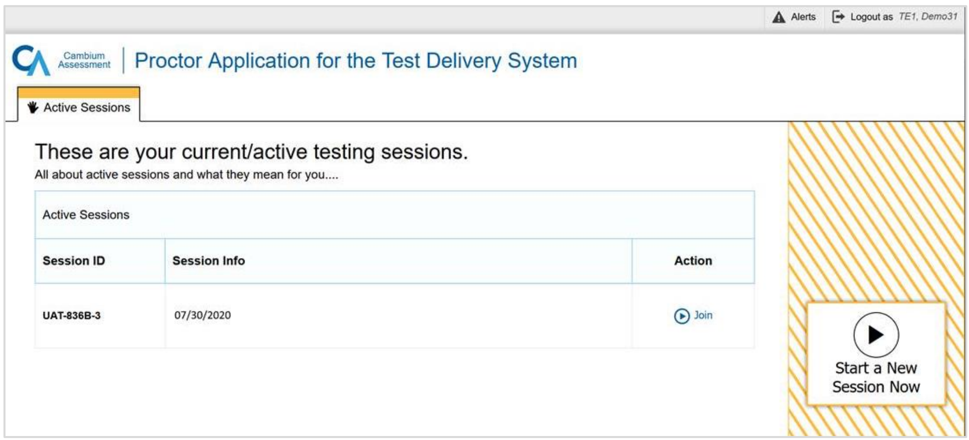
You can transfer an active test session from one device or browser to another without stopping the session or interrupting in-progress tests. This is useful if your computer malfunctions or if you accidentally close the browser while a session is in progress.

Your session remains open until it times out. If you do not return to the active session within 20 minutes and there is no student activity during that time, the Online Testing System logs you out and pauses the students' tests.

The Online Testing System ensures that you can administer a test session only from one browser at a time. If you move a test session to a new device, you cannot simultaneously administer the session from the original browser or device.

1. While the session is still active on the original device or browser, log in to the TA Site on the new device or browser. The ***Active Sessions*** page appears (see Figure 61), listing the active session.

Figure 61. Active Sessions Page



2. Select **Join**. The test session page appears, allowing you to continue monitoring your students' progress. The test session on the previous computer or browser closes automatically. If the test session displays a timer, the timer continues counting down where it left off.

If you do not wish to return to the active session, you can select **Start a New Session Now** to open the ***Test Selection*** window and create a new test session.

User Support and Troubleshooting Information

For information and assistance in using the Online Testing System, contact the Montana Program Help Desk. The Help Desk is open Monday–Friday 6:00 a.m. to 6:00 p.m. (except holidays or as otherwise indicated on the Montana Testing portal).

Montana Help Desk

Toll-Free Phone Support: 877-365-7915

Email Support: mthelpdesk@cambiumassessment.com

Please provide the Help Desk with a detailed description of your problem, as well as the following:

- If the issue pertains to a student, provide the associated district or school for that student. SSID and additional student PII can only be communicated via phone with the help desk**. Do not leave any student identifying information such as a student name, SSID, and/or personal characteristics in a voicemail or email.**
- Test Administrator name.
- If the issue pertains to a TIDE user, provide the user's full name and email address.
- Any error messages and codes that appeared, if applicable.
- Affected test ID and question number, if applicable.
- Operating system and browser version information, including version numbers (for example, Windows 10 and Firefox 87).
- Information about your network configuration, if known:
  - Secure Browser installation (to individual devices or network).
  - Wired or wireless internet network setup.

Troubleshooting

This section provides troubleshooting tips for common issues that may occur while testing.

Username and Password Issues

Your username for logging in to the TA Site is the email address associated with your account in TIDE. When you are added to TIDE, you receive an activation email containing a temporary link to the ***Reset Your Password*** page. To activate your account, you must set up your password within 15 minutes of the email being sent.

- **If your first temporary link expired:**

  In the activation email you received, select the second link provided and request a new temporary link.

- **If you forgot your password:**

  On the ***Login*** page, select **Forgot Your Password?** and then enter your email address in the *Email Address* field. You will receive an email with a new temporary link to reset your password.

- **If you did not receive an email containing a temporary link or authentication code:**

  Check your spam folder to make sure your email program did not categorize it as junk mail. If you still do not have an email, contact your School or District Test Coordinator to make sure you are listed in TIDE.

- **Additional help:**

  If you are unable to log in, contact the State Assessment Program Help Desk for assistance. You must provide your name and email address. Contact information is available in the User Support and Troubleshooting Information section.

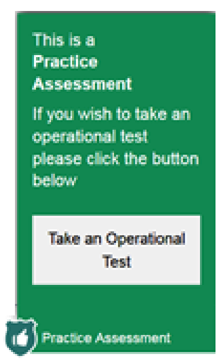
Common Student Sign-in Errors

The Online Testing System generates an error message if a student cannot sign in. The following are the most common student sign-in issues:

- **Session does not exist:**

  The student entered the Session ID incorrectly or signed in to the wrong site. Verify that the student correctly entered the active Session ID. Also, verify that both you and the student are using the correct sites. For example, students signed in to the Student Training Site cannot access sessions created in the TA Interface. A message in the bottom-left corner of the ***Student Sign-In*** page indicates which site the student is on. If a student is on the wrong site, the student can select the button in the message to switch to the correct site.

Figure 62. Student Training Site Badge on the Login Page



- **Student information is not entered correctly:**

  Verify that the student entered the SSID correctly. If this does not resolve the error, use the Student Lookup tool to verify the student's information.

- **Session has expired:**

  The Session ID corresponds to a closed session. Ensure that the student enters the correct Session ID and verify that your session is open. For more information about test sessions, see the section How to Select Tests and Start a Test Session.

- **Student is not associated with the school:**

  The student is not associated with your school, or you are not associated with the student's school.

Resolving Secure Browser Error Messages

This section provides possible resolutions for the following messages that students may receive when signing in to tests using the Secure Browser.

- **You cannot login with this browser:**

  This message occurs when the student is not using the correct Secure Browser. To resolve this issue, ensure the latest version of the Secure Browser is installed, and that the student launched the Secure Browser instead of a standard web browser. If the latest version of the Secure Browser is already running, then log out the student, restart the device, and try again.

- **Looking for an internet connection…:**

  This message appears when the Secure Browser cannot connect with the Online Testing System. This can occur if there is a network-related problem. Make sure that either the network cable is plugged in (for wired connections) or the Wi-Fi connection is live (for wireless connections). Also check if the Secure Browser must use specific proxy settings; if so, those settings must be specified as options when configuring the Secure Browser. If connection issues persist, contact a network technician.

- **Test Environment Is Not Secure:**

  This message can occur when the Secure Browser detects a denylisted application running on the device. If this message appears on an iPad, ensure that either Assessment Mode is enabled.

Change Log

Location	Change	Date
Throughout	Clarifying language and screenshot replacements throughout. Additional clarifying language for search/filter feature in TA interface.	8/11/22

Appendix 5-F
Usability, Accessibility, and Accommodations Guidelines

Montana Comprehensive Assessment System

Usability, Accessibility, and Accommodations Guidelines

Published August 11, 2022

Montana Science Assessments

Table of Contents

Table of Tables

INTRODUCTION

This document is intended to provide guidance for Montana school district personnel who must make decisions about testing all students, including special student populations, on the Montana Science Assessment. All students, including students with disabilities and English Learners, are required by state and federal law to take part in the state assessments with or without accommodations. Generally, if a student can receive instruction, then they are also able to participate in state assessments. All students enrolled in accredited schools are expected to take part in state assessments in one of three ways:

1. Participate in general population assessments **without** accommodations.
2. Participate in general population assessments **with** accommodations.
3. Participate in **alternate assessments** when the participation criteria are met.

Special student populations include students identified as eligible for special education services under the Individuals with Disabilities Education Act (IDEA), students identified as disabled under Section 504 of the Rehabilitation Act of 1973, and students who are identified as English learners (EL). This document also provides information regarding general education students who may require additional supports when taking standardized assessments.

These *Guidelines* describe the universal tools, designated supports, and accommodations available for the following assessments:

- Montana Science Interim Assessments
- Montana Science Summative Assessments

The guidelines available for special student populations eligible to participate in the the Alternate Montana Science Assessment (AMSA) and can be found on the Montana Testing Portal under the Accessibility and Accommodations resources.

The specific accessibility features (i.e., universal tools, designated supports, and accommodations) approved by the Montana Office of Public Instruction (OPI) are subject to change in the future if additional tools, supports or accommodations are identified for the assessments based on experience and research findings.

Student eligibility for accessibility features is added through the state student information system known as Achievement In Montana (AIM)/Infinite Campus by District-Level users. The AIM/Infinite Campus system is designed to collect demographic, enrollment, program participation, and assessment data for each student.

The OPI uploads the student data file from AIM/Infinite Campus into the Test Information Distribution Engine (TIDE) (see TIDE OPI Student File Upload Schedule). TIDE then distributes this information to the appropriate system**.** To protect student data privacy, districts are responsible for turning on any accessibility tools within the Montana Testing Portal. Specific data within an IEP is not transferred from AIM into TIDE. In addition to this document, OPI continues to prepare teachers to administer Montana Science Assessment to all students. Resources related to student supports and accessibility needs are available on the Montana Testing Portal.

Intended Audience and Recommended Use

This document is intended for school-level personnel and decision-making teams, particularly Individualized Education Plan (IEP) teams, as they prepare for and implement the Montana Science Assessment. The *Guidelines* provide information for classroom teachers, English development educators, special education teachers, and related services personnel to use in selecting and administering universal tools, designated supports, and accommodations for those students who need them. The *Guidelines* are also intended for assessment staff and administrators who oversee the decisions that are made in instruction and assessment.

The *Guidelines* apply to **all** students. They emphasize an individualized approach to the implementation of assessment practices for those students who have diverse needs and participate in large-scale content assessments. This document focuses on universal tools, designated supports, and accommodations for the Montana Science Assessment. It recognizes and supports the critical connection between accessibility and accommodations in instruction and during assessment.

The selection of accessibility supports (universal tools, designated supports, and accommodations) is a systematic, data-based, needs-based process that is made by educators familiar with individual student needs. For students being served with English learner plans, IEPs, or 504 plans, supports should be discussed at plan meetings. For at-risk students who would benefit from designated supports, a team of educators familiar with the student should make accessibility support decisions. The supports that are provided on state assessments must be familiar to the student and match those supports and accommodations that are provided for classroom instruction and assessments throughout the school year. For additional guidance, please refer to Montana's Three Tiers of Accessibility.

Figure 1. Conceptual Model of Montana Accessibility Features

Note: Not all tools are available on all assessments. Review each section and tool carefully to determine if it is applicable to the assessment being administered. If you have any questions regarding tool availability, refer to the OPI Assessment Help Desk at 1-844-867-2569 or OPIAssessmentHelpDesk@mt.gov.

The Conceptual Model recognizes that all students should be held to the same expectations for instruction in Montana Content Standards and have available to them universal accessibility features. It also recognizes that some students may have certain characteristics and access needs that require the use of accommodations for instruction when they participate in the assessments.

These *Guidelines* present the current universal tools, designated supports, and accommodations adopted by the OPI to ensure valid assessment results for all students taking state mandated assessments.

SECTION I: UNIVERSAL TOOLS

WHAT ARE UNIVERSAL TOOLS?

Universal tools are accessibility resources of the assessment that are either provided as digitally delivered components of the test administration system or separate from it. Universal tools are available to all students based on student preference and selection. The universal tools described in this section are not modifications and all yield valid scores that count as participation in assessments that meet the requirements of the Every Student Succeeds Act (ESSA) when used in a manner consistent with the *Guidelines*.

EMBEDDED UNIVERSAL TOOLS

Digitally delivered assessments include a wide array of embedded universal tools. These are available to all students as part of the technology platform. Table 1 lists the embedded universal tools available to all students for Montana Science Assessment computer-administered assessments. It includes a description of each tool. Although these tools are available to all students, educators may determine that one or more might be distracting for a particular student, and thus might indicate that the tool should be turned off for the administration of the assessment to the student (see Designated Supports). Universal Tools are turned enabled by default but can be turned OFF locally in the Test Administrator Interface prior to administering tests or by the student. Note: Universal Tools cannot be set in the TIDE > Test Settings. Test Administrators (TA) and Proctors can check a student's test settings in the Test Settings screen in the TA Interface and update them as required.

Table 1. Embedded Universal Tools Available to All Students

Universal Tool	Description
Breaks (Embedded)	The number of items per session can be flexible based on the student's need. Breaks of more than 20 minutes will prevent the student from returning to items already attempted by the student. There is no limit on the number of breaks that a student might be given. The use of this universal tool may result in the student needing additional overall time to complete the assessment. Refer to the MSA Test Administration Manual and TIDE User Guide for additional guidance.
Calculator (Embedded) *(See non-embedded accommodations for students who cannot use the embedded calculator)*	An embedded on-screen digital calculator can be accessed when students click on the calculator button. When the embedded calculator, as presented for all students, is not appropriate for a student (for example, for a student who is blind), the student may use the calculator offered with assistive technology devices (such as a talking calculator or a braille calculator).
Digital Notepad	This tool is used for taking notes about an item. The digital notepad is item-specific and is available through the end of the test segment. Notes are not saved when the student moves on to the next segment or after a break of more than 20 minutes.

Universal Tool	Description
Expandable Passages/Items	Each passage and/or associated item(s) can be expanded so that it takes up a larger portion of the screen. This can be accessed within the testing environment by pressing the (← →) button on the screen.
Global Notes	Global notes is a notepad feature that is available throughout the entirety of the test. The student clicks on the notepad icon for the notepad to appear. The notes are retained after breaks or pauses in the test.
Highlighter (Embedded)	A digital tool for marking desired text, item questions, item answers, or parts of these with a color. Highlighted text remains available throughout each test segment.
Keyboard Navigation (Embedded)	Navigation throughout text can be accomplished by using a keyboard.
Line Reader (Embedded)	The student uses an onscreen universal tool to assist in reading by raising and lowering the tool for each line of text on the screen.
Mark for Review (Embedded)	Allows students to flag items for future review during the assessment. Markings are not saved when the student moves on to the next segment or after a break of more than 20 minutes.
Periodic Table	An arrangement of the chemical elements, ordered by their atomic number, electron configuration, and recurring chemical properties. The ordering shows periodic trends, such as elements with similar behavior in the same column. For students testing with Spanish translations as a designated support, this feature will be available in Spanish language.
Strikethrough	Allows users to cross out answer options. If an answer option is an image, a strikethrough line will not appear, but the image will be grayed out.

Non-embedded Universal Tools

Some universal tools may need to be provided outside of the computer test administration system. These tools, shown in Table 2, are to be provided locally for students. They can be made available to any student.

Table 2. Non-embedded Universal Tools Available to All Students

Universal Tool	Description
Breaks (Non-Embedded)	Breaks may be given at predetermined intervals or after completion of sections of the assessment for students taking a paper-based test. Sometimes students are allowed to take breaks when individually needed to reduce cognitive fatigue when they experience heavy assessment demands. The use of this universal tool may result in the student needing additional overall time to complete the assessment.
Calculator (Non-Embedded)	When the embedded calculator tool is not suitable for a student participating in the assessment, the provision of a battery-operated hand-held calculator may be appropriate. If a calculator is provided, proctors must ensure that the device is functional, has working batteries, and that the student is familiar and comfortable with how to use it. Students may not use calculators available on their phones, iPads, or other electronic devices. Students may not share calculators.
Scratch Paper	Students may use blank scratch paper to make notes, write computations, record responses, or create graphic organizers. A whiteboard with a marker may be used as scratch paper. As long as the construct being measured is not impacted, assistive technology devices, including low-tech assistive technology (Math Window), are permitted to make notes, including the use of digital graph paper. The assistive technology device needs to be familiar to the student and/or consistent with the child's IEP or 504 plan. Access to internet must be disabled on assistive technology devices. If a student needs to take the assessment in more than one session, scratch paper, whiteboards, and/or assistive technology devices may be collected at the end of each session, securely stored, and made available to the student at the next testing session. Once the student completes the test, the scratch paper must be collected and securely destroyed, whiteboards should be erased, and notes on assistive technology devices erased to maintain test security.
Thesaurus	A thesaurus contains synonyms of terms while a student interacts with text included in the assessment. The use of this universal tool may result in the student needing additional overall time to complete the assessment.

SECTION II: DESIGNATED SUPPORTS

What Are Designated Supports?

Designated supports for the Montana Science Assessment are those features that are available for use by **any student** for whom the need has been indicated by an educator (or team of educators) with the parent/guardian and student. The designated supports described in this section are not modifications. Designated supports all yield valid scores that count as participation in assessments that meet the requirements of ESSA when used in a manner consistent with the *Guidelines*. It is recommended that a consistent process be used to determine these supports for individual students. All educators making these decisions should be trained on the process and should be made aware of the range of available designated supports. The OPI has identified digitally embedded and non-embedded designated supports for students for whom an adult or team has indicated a need for the support.
Designated supports need to be identified prior to assessment administration. Embedded and non-embedded supports must be entered into TIDE under student test settings as no accessibility features automatically transfer from AIM/Infinite Campus. Any non-embedded designated supports must be arranged for prior to testing and provided during testing.

Who Makes Decisions About Designated Supports?

Informed adults make decisions about designated supports. Ideally, the decisions are made by all educators familiar with the student's characteristics and needs, as well as those supports that the student has been using during instruction and for other assessments. Student and parent input to the decision is also recommended.

Embedded Designated Supports

Table 3 lists the embedded designated supports available to all students for whom the need has been indicated. It includes a description of each support along with recommendations for when the support might be needed.

Table 3. Embedded Designated Supports

Designated Support	Description	Recommendations for Use
Color Choices	Enable students to adjust screen background or font color, based on student needs or preferences. This may include reversing the colors for the entire interface or choosing the color of font and background. Available options include: Blue, Light blue, Black on cream, Gray, Light gray, Medium Gray on Light Gray Green, Light green, Magenta, Light magenta, White on navy, White on red, Red on white, Yellow, Light yellow,	Students with attention difficulties may need this support for viewing test content. It also may be needed by some students with visual impairments or other print disabilities (including learning disabilities). Choice of colors should be informed by evidence that color selections meet the student's needs.

Designated Support	Description	Recommendations for Use
	Yellow on Blue; Yellow on black, and Reverse Contrast.	
Masking	Masking involves blocking off content that is not of immediate need or that may be distracting to the student. Students are able to focus their attention on a specific part of a test item by masking.	Students with attention difficulties may need to mask content not of immediate need or that may be distracting during the assessment. This support also may be needed by students with print disabilities (including learning disabilities) or visual impairments. Masking allows students to hide and reveal individual answer options, as well as all navigational buttons and menus.
Mouse Pointer (Size and Color)	This embedded support allows the mouse pointer to be set to a larger size and also for the color to be changed. A test administrator sets the size and color of the mouse pointer prior to testing.	Students who are visually impaired and need additional enlargement or a mouse pointer in a different color to find their mouse pointer more readily on the screen will benefit from the mouse pointer support. Students who have visual perception challenges will also find this beneficial. The size and color are set during registration and cannot be changed during the administration of the assessment. Students should have ample opportunity to practice during daily instruction with the size and color to determine student preference. The mouse pointer can be used with the zoom designated support. If students are using a magnification program (See Designated Support, magnification), the enlarged mouse pointer is built into magnification programs and mouse pointer may not be needed. It is recommended that students requiring this support test on a device with an external mouse including a scroll wheel.
Permissive Mode	This support allows approved 3rd party AT software to be	The secure browser is designed to purposefully block access to most additional software to

Designated Support	Description	Recommendations for Use
	accessible while testing via the secure browser.	create a secure testing environment. In some cases, students may need access to an approved 3rd party software to appropriately interact with the test content according to their IEP/504 plan. Permissive Mode may be applied to allow access to appropriate AT software during testing.
Streamlined Mode	This designated support provides a streamlined interface of the test in an alternate, simplified format in which the items are displayed below the stimuli.	This designated support may benefit a small number of students who have specific learning and/or reading disabilities and/or visual impairment in which the text is presented in a more sequential format. Students should have familiarity interacting with items in streamline format.
Text-to-Speech *Text-to-Speech (Items and Stimuli)* *Text-to-Speech (Items only)* *Text-to-Speech (Stimuli only)*	Text is read aloud to the student via embedded text-to-speech technology. The student is able to control the speed as well as raise or lower the volume of the voice via a volume control. This feature may be applied to all test content (Items and Stimuli), Items Only, or Stimuli Only, as appropriate.	Students who are struggling readers may need assistance accessing the assessment by having all or portions of the assessment read aloud. This support also may be needed by students with reading-related disabilities, or by students who are blind and do not yet have adequate braille skills. Students would need to use this support regularly during instruction to meaningfully benefit from it on assessments. Students who use text-to-speech will need headphones unless tested individually in a separate setting.
Translations (Toggle)	This Spanish presentation will allow the literate Spanish-speaking student to toggle between a full Spanish translation of the item and the English version of the item. By default, all test directions, navigation buttons, and test content will be presented to the student in the Spanish language.	For students whose primary language is not English and who use dual language supports in the classroom, use of the stacked (dual language) translation may be appropriate. Students participate in the assessment regardless of the language. This support will increase reading load and cognitive load. The use of this support may result in the student

Designated Support	Description	Recommendations for Use
		needing additional overall time to complete the assessment.
Translated Test Directions (Spanish)	Translation of test directions is a language support available prior to beginning the actual test items. Students can see test directions in another language. As an embedded designated support, translated test directions are automatically a part of the Spanish translations designated support.	Students who are advancing toward English language proficiency (including non-ELs, ELs, and ELs with disabilities) can use the translated directions support. This support should only be used for students who are proficient readers in the other language and not proficient in English.
Turn Off Any Universal Tools	Disabling any universal tools that might be distracting or that students do not need to use or are unable to use.	Students who are easily distracted (whether or not designated as having attention difficulties or disabilities) may be overwhelmed by some of the universal tools. Knowing which specific tools may be distracting is important for determining which tools to turn off.
Zoom	Setting the default text size or other graphics in a window or frame to appear larger on the screen. The default font size for all tests is 14 pt. To increase the default print size of the entire test, the print size must be set for the student in TIDE or set by the test administrator prior to the start of the test. The use of this tool may result in the student needing additional overall time to complete the assessment. Students can still have the ability to alter the print size on individual test pages, by utilizing the *Zoom In* and *Zoom Out* buttons as a universal tool.	Students used to viewing enlarged text or graphics, or navigation buttons, with or without changes to color contrast, may need zoom to comfortably view content. This support also may meet the needs of students with visual impairments and other print disabilities. The use of this designated support may result in the student needing additional overall time to complete the assessment.

Non-embedded Designated Supports

Some designated supports may need to be provided outside of the digital-delivery system. These supports, shown in Table 4, are to be provided locally for those students unable to use the designated supports when provided digitally.

Table 4. Non-embedded Designated Supports

Designated Support	Description	Recommendations for Use
Amplification	The student adjusts the volume control beyond the computer's built in settings using headphones or other non-embedded devices.	Students may use amplification assistive technology (e.g., headphones, FM System, noise buffers, white noise machines) to increase the volume provided in the assessment platform. Use of this resource likely requires a separate setting. If the device has additional features that may compromise the validity of the test (e.g., internet access), the additional functionality must be deactivated to maintain test security.
Bilingual Dictionary	A bilingual/dual language word-to-word dictionary is a language support. A bilingual/dual language word-to-word dictionary can be provided.	For students whose primary language is not English and who use dual language supports in the classroom, use of a bilingual/dual language word-to-word dictionary may be appropriate. Students participate in the assessment regardless of the language. The use of this support may result in the student needing additional overall time to complete the assessment.
Color Contrast	Test content of online items may be printed with different colors.	Students with attention difficulties may need this support for viewing the test when digitally provided color contrasts do not meet their needs. Some students with visual impairments or other print disabilities (including learning disabilities) also may need this support. Choice of colors should be informed by evidence of those colors that meet the student's needs.
Color Overlay	Color transparencies are placed over a paper-based assessment.	Students with attention difficulties may need this support to view test content. This support also may be needed by some students with visual impairments or other print disabilities (including learning disabilities). Choice of color should be informed by evidence of those colors that meet the student's needs.

Designated Support	Description	Recommendations for Use
Magnification (Non-Embedded)	The size of specific areas of the screen (e.g., text, formulas, tables, graphics, navigation buttons, and mouse pointer) may be adjusted by the student with an assistive technology device or software. Magnification allows increasing the size and changing of the color contrast, including the size and color of the mouse pointer, to a level not provided for by the zoom universal tool, color contrast designated support, and/or mouse pointer designated support.	Students used to viewing enlarged text or graphics, or navigation buttons, with or without changes to color contrast, may need magnification to comfortably view content. This support also may meet the needs of students with visual impairments and other print disabilities. The use of this designated support may result in the student needing additional overall time to complete the assessment.
Medical Supports	Students may have access to medical supports for medical purposes (e.g., Glucose Monitor). The medical support may include a cell phone and should only support the student during testing for medical reasons.	Educators should follow local policies regarding medical supports and ensure students' health is the highest priority. Electronic medical support settings must restrict access to other applications or the test administrator must closely monitor the use of the medical support to maintain test security. Use of medical supports may require a separate setting to avoid distractions to other test takers and to ensure test security.
Noise Buffers	Ear mufflers, white noise, and/or other equipment used to block external sounds.	Student (not groups of students) wears equipment to reduce environmental noises. Students may have these testing variations if regularly used in the classroom. Students who use noise buffers will need headphones unless tested individually in a separate setting.
Read Aloud Items/Stimuli	Text is read aloud to the student by a trained and qualified human reader who follows the administration guidelines provided in the *Test Administration Manual* and *Read Aloud Guidelines* (see Appendix B). All or portions of the content may be read aloud.	Students who are struggling readers may need assistance accessing the assessment by having all or portions of the assessment read aloud. This support also may be needed by students with reading-related disabilities, or by students who are blind and do not yet have adequate braille skills. If not used regularly during instruction, this support is likely to be confusing and may impede the performance on assessments. Readers should be provided to students on an individual basis – not to a group of students. A student should have the option of asking a reader to slow down or

Designated Support	Description	Recommendations for Use
		repeat text. The use of this support may result in the student needing additional overall time to complete the assessment and/or the use of a separate setting.
Read Aloud – Spanish Items/Stimuli	Spanish text is read aloud to the student by a trained and qualified human reader who follows the administration guidelines provided in the *Test Administration Manual* and the *Read Aloud Guidelines*. All or portions of the content may be read aloud.	Students receiving the translations (stacked) designated support and who are struggling readers may need assistance accessing the assessment by having all or portions of the assessment read aloud. This support also may be needed by students with reading-related disabilities. If not used regularly during instruction, this support is likely to be confusing and may impede the performance on assessments. A student should have the option of asking a reader to slow down or repeat text. The use of this support may result in the student needing additional overall time to complete the assessment and/or the use of a separate setting.
Scribe Items (Non-Writing)	Students dictate their responses to a human who records verbatim what they dictate. The scribe must be trained and qualified and must follow the administration guidelines provided in Appendix C: Scribing Guidelines.	Students who have documented significant motor or processing difficulties, or who have had a recent injury (such as a broken hand or arm) that make it difficult to produce responses may need to dictate their responses to a human, who then records the students' responses verbatim. The use of this support may result in the student needing additional overall time to complete the assessment.
Separate Setting	Test location is altered so that the student is tested in a setting different from that made available for most students.	Students who are easily distracted (or may distract others) in the presence of other students, for example, may need an alternate location to be able to take the assessment. The separate setting may be in a different room that allows them to work individually or among a smaller group. The student may read aloud to self, use a device requiring voicing (e.g., a Whisper Phone), or use Amplification. It may also include a calming device or support as recommended by educators and/or specialists. Or, the separate setting may be in the same room but in a specific location (for example, away from windows, doors, or pencil sharpeners, in a study carrel, near the teacher's desk, or in the front of a classroom). Some students may benefit from being in an environment

Designated Support	Description	Recommendations for Use
		that allows for movement, such as being able to walk around. In some instances, students may need to interact with instructional or test content outside of school, such as in a hospital or their home. A specific adult, trained in a manner consistent with the Test Administrator Manual (TAM), can act as test proctor when student requires it.
Simplified Test Directions	The test administrator simplifies or paraphrases the test directions found in the *Test Administration Manual*.	Students who need additional support understanding the test direction may benefit from this resource. This designated support may require testing in a separate setting to avoid distracting other test takers.
Translated Test Directions (Non-Embedded)	PDF of directions translated in each of the languages currently supported. Bilingual adult can read to student. Note: Translated Test Directions are currently only available in Spanish.	Students who have limited English language skills (whether or not designated as ELs or ELs with disabilities) can use the translated test directions. In addition, a biliterate adult trained in the test administration manual can read the test directions to the student. The use of this support may result in the student needing additional overall time to complete the assessment.

SECTION III: ACCOMMODATIONS

What Are Accommodations?

Accommodations are changes in procedures or materials that increase equitable access during the Montana Science Assessment. The accommodations described in this section are not modifications. Accommodations all yield valid scores that count as participation in assessments that meet the requirements of ESSA when used in a manner consistent with the *Guidelines*. They allow students to show what they know and can do. The OPI has identified digitally-embedded and non-embedded accommodations for students with disabilities. Accommodations must be documented in an Individualized Education Plan (IEP) or 504 Plan. If an accommodation is not visible to select from or an emergency need has risen, requests for **Non-routine accommodations** must be submitted to the OPI from the System Test Coordinator (STC) through the MontCAS Application > Non-Routine Request Process. The OPI must approve all non-routine accommodations used in the state assessments **before** a student can receive them as any accommodation not listed has the risk to change what is being measured and may make the student score invalid.

Accommodations help certain students access the general education curriculum and should be used in instruction and assessment throughout the year and at least 3 months before the assessment. Supports that are provided on state assessments must be familiar to the student and must match those supports and accommodations that are provided for classroom instruction and assessments throughout the school year. Student's parents/guardians must be knowledgeable about the supports and accommodations planned for their child so they are aware of the conditions under which their child will participate in the assessment.

Who Makes Decisions About Accommodations?

Decisions about accommodations are made by the IEP teams and educators. These teams (or educators for 504 plans) provide evidence of the need for accommodations and ensure they are noted on the IEP or 504 plan for students with disabilities who require accommodations. Members of these teams always include the parent and/or guardian of the student. Individuals with Disabilities Education Act (IDEA) requires specific members of the IEP team. These may include the student, an administrator, special education teachers, related service providers, a school psychologist, and general education teachers of the student or teachers with grade level content knowledge.

The IEP team (or educator developing the 504 plan) is responsible for ensuring that the IEP is created in the local Student Information System (SIS), which syncs into AIM/Infinite Campus. The district's STC or designated user roles with permissions above "BC" (see User Roles and Access Document) is responsible for making sure that the accommodations and supports that are in the IEP or the 504 plan are entered into TIDE in the Test Settings module. A student without a documented IEP/504 in AIM/Infinite Campus is unable to have any accommodations turned on for them in the Montana Testing Portal [i.e., TIDE]. All embedded accommodations must be activated prior to testing. Note: accessibility features do not automatically transfer from AIM/infinite Campus into TIDE. This is a school district responsibility to ensure the test settings are properly configured on an annual basis per each child's education plan. Any embedded test setting modification MUST be present in the TIDE test settings module in order to render within the student testing interface.

Determination of which accommodations an individual student will have available for them to meaningfully participate in must be determined before the assessment and locally turned ON in the TIDE > Test Setting module to enable embedded accommodations. Students requiring non-embedded accommodations must also have these determinations made in advance of participating in the assessment.

Embedded Accommodations

Table 5 lists the embedded accommodations available for the Montana Science Assessment for those students for whom the accommodations are included on an IEP or 504 plan. The table includes a description of each accommodation along with recommendations for when the accommodation might be needed and how it can be used. For those accommodations that may be considered controversial, a description of considerations about the use of the accommodation is provided.

Table 5. Embedded Accommodations

Accommodation	Description	Recommendations for Use
Braille (Embedded)	A raised-dot code that individuals read with the fingertips. Graphic material (e.g., maps, charts, graphs, diagrams, and illustrations) is presented in a raised format (paper or thermoform). Nemeth and UEB contracted Technical code(s) are available for science.	Students with visual impairments may read text via braille. Tactile overlays and graphics also may be used to assist the student in accessing content through touch. Due to limitations with refreshable braille technology, some braille items and graphics will be presented via embosser. Alternative text descriptions are embedded in the assessment for all graphics. The use of this accommodation may result in the student needing additional overall time to complete the assessment.

Non-embedded Accommodations

Table 6 lists the non-embedded accommodations available for the Montana Science Assessment for those students for whom the accommodations are documented on an IEP or 504 plan. The table includes a description of each accommodation, along with recommendations for when the accommodation might be needed and how it can be used.

Table 6. Non-embedded Accommodations

Accommodation	Description	Recommendations for Use
Alternate Response Options	Alternate response options include but are not limited to adapted keyboards, large keyboards, Sticky Keys, Mouse Keys, FilterKeys, adapted mouse, touch screen, head wand, and switches.	Students with some physical disabilities (including both fine motor and gross motor skills) may need to use the alternate response options accommodation. Some alternate response options are external devices that must be plugged in and be compatible with the assessment delivery platform.
Braille (Non-Embedded)	A raised-dot code that individuals read with the fingertips. Graphic material (e.g., maps, charts, graphs, diagrams, and illustrations) is presented in a raised format (paper or thermoform). A brailled paper version of the test is available as an option for administration, as needed.	Students with visual impairments may read text via braille. Tactile overlays and graphics also may be used to assist the student in accessing content through touch. The use of this accommodation may result in the student needing additional overall time to complete the assessment.
Print on Request *Print on Request (Items)* *Print on Request (Stimuli)* *Print on Request (Stimuli and Items)*	Paper copies of either passages/stimuli and/or items are printed for students. For those students needing a paper copy of a passage or stimulus, permission for the students to request printing must first be set in TIDE.	Some students with disabilities may need paper copies of either passages/stimuli and/or items. A very small percentage of students should need this accommodation. The use of this accommodation may result in the student needing additional time to complete the assessment. Please note that Print on Request is the only printed form of accommodation available for students needing paper copy of items.
Read Aloud Passages	Text is read aloud to the student via an external screen reader or by a trained and qualified human reader who follows the administration guidelines provided in the *Read Aloud Guidelines* (See	This accommodation is appropriate for a very small number of students. Read aloud is available as an accommodation for students whose need is documented in an IEP or 504 plan. A student should have the option of asking a reader to slow down or repeat text. The use of this

Accommodation	Description	Recommendations for Use
	Appendix B) All or portions of the content may be read aloud.	accommodation may result in the student needing additional time to complete the assessment and/or the use of a separate setting.
Sign Language – Test Items	Test content may be translated into Sign Language by a human signer.	Some students who are deaf or hard of hearing and who typically use Sign Language may need this accommodation when accessing text-based content in the assessment. The use of this accommodation may result in the student needing additional overall time to complete the assessment. For many students who are deaf or hard of hearing, viewing signs is the only way to access information presented orally.
Specialized Calculator	A non-embedded, stand-alone calculator for students needing a specialized calculator, such as a braille calculator or a talking calculator, currently unavailable within the assessment platform.	Students who are unable to use the embedded calculator for calculator-allowed items will be able to use the calculator that they typically use, such as a braille calculator or a talking calculator. Test administrators should ensure that the calculator is available only for designated calculator items.
Speech-to-Text (Non-Embedded)	Voice recognition allows students to use their voices as input devices to the computer, to dictate responses or give commands (e.g., opening application programs, pulling down menus, and saving work). Voice recognition software generally can recognize speech up to 160 words per minute. Students may use their own assistive technology devices.	Students who have motor or processing disabilities (such as dyslexia) or who have had a recent injury (such as a broken hand or arm) that make it difficult to produce text or commands using computer keys may need alternative ways to work with computers. Students will need to be familiar with the software and have had many opportunities to use it prior to testing. Speech-to-text software requires that the student go back through all generated text to correct errors in transcription, including use of writing conventions; thus, prior experience with this accommodation is essential. If students use their own assistive technology devices, all assessment content should be deleted from these devices after the test for security purposes. For many of

Accommodation	Description	Recommendations for Use
		these students, using voice recognition software is the only way to demonstrate their composition skills. Still, use of speech-to-text does require that students know writing conventions and that they have the review and editing skills required of students who enter text via the computer keyboard. It is important that students who use speech-to-text also be able to develop planning notes via speech-to-text, and to view what they produce while composing via speech-to-text.
Timing or Scheduling	Students can be tested during their optimal time of day. Scheduling should account for a student who requires their test to be frequently paused over an extended time period.	The Montana Science Assessment is not a timed assessment. Students may take as many breaks as needed throughout the assessment as determined by the student's instructional team.

APPENDIX A. SUMMARY OF TOOL DESIGNATION FOR STUDENTS

	All Students	English learners (ELs)	Students with disabilities	ELs with disabilities
Universal Tools	✓	✓	✓	✓
Designated Supports	✓[1]	✓[1]	✓	✓
Accommodations			✓	✓

[1] *Only for instances that an adult (or team) has deemed the supports appropriate for a specific student's testing needs.*

APPENDIX B. READ ALOUD GUIDELINES

When a student cannot access text-to-speech, which is an embedded resource available on Montana's Testing Portal, the student may be eligible to work with a test reader. A test reader is an adult who provides an oral presentation of the assessment text to an eligible student. The student depends on the test reader to read the test questions accurately, pronounce words correctly, and speak in a clear voice throughout the test. The test reader must be trained and qualified and must follow the *Read Aloud Guidelines* presented here. The guiding principle in reading aloud is to ensure that the student has access to test content.

Test readers are allowable across all grades as a **designated support** for science stimuli and items. Note that this accommodation is appropriate for a very small number of students (estimated to be approximately 1-2% of students with disabilities participating in a general assessment).

QUALIFICATIONS FOR TEST READERS

- The test reader should be an adult who is familiar with the student and who is typically responsible for providing this support during educational instruction and assessments.
- Test readers must be trained on the administration of the assessment in accordance with Montana Office of Public Instruction policy, and familiar with the terminology and symbols specific to the test content and related conventions for standard oral communication.
- Test readers must be trained in accordance with Montana Testing Portal administration, as well as security policies and procedures as articulated in the Montana Science Assessment test administration manuals, guidelines, and related documentation.

PREPARATION

- Test readers should read and sign a test security/confidentiality agreement prior to test administration. If test readers have a TIDE account and have taken the TA Certification Course, then their Test Security Agreement (TSA) will have been signed online. If not, the System Test Coordinator (STC) should ensure that the test reader signs a paper copy of the TSA.
- Test readers are expected to familiarize themselves with the test environment and format in advance of the testing session. Having a working familiarity with the test environment and format will help facilitate reading of the test.
- Test readers should have a strong working knowledge of the embedded and non-embedded universal tools, designated supports, and accommodations available on Montana 's Testing Portal's online assessments.
- Test readers should be aware of and familiar with all additional supports and/or accommodations provided to a student in accordance with the student's Individualized Education Program (IEP) or 504 plan. This will ensure that there are plans in place for providing all needed designated supports and accommodations.

- In addition to a test reader, students may make use of any other approved specialized tools or equipment during the test as appropriate and in accordance with the *Usability, Accessibility, and Accommodations Guidelines*. Test readers should be familiar with any assistive technology or approved supports the student requires.
- Test readers should have extensive practice in providing read aloud support and must be familiar and comfortable with the process before working directly with a student.
- The reader should be knowledgeable of procedures for reading aloud text by content area (see Table 7 below).
- The test reader should meet with the student in advance and inform the student of the parameters of the support. A suggested test reader script is included at the end of the *Read Aloud Guidelines*.
- Unless otherwise specified by a student's IEP or 504 plan, the test reader does not have a role in manipulating the test or assisting with any other support tools. Test readers should be ready with the appropriate script that reinforces the parameters during the test session.

GENERAL GUIDELINES

- The test reader's support should ideally be provided in a separate setting so as not to interfere with the instruction or assessment of other students.
- Read each question exactly as written and as clearly as possible.
- Throughout the exam, strive to communicate in a neutral tone and maintain a neutral facial expression and posture.
- Avoid gesturing, head movements, or any verbal or non-verbal emphasis on words not otherwise emphasized in text.
- Avoid conversing with the student about test questions as this would be a violation of test security; respond to the student's questions by repeating the item, words or instructions verbatim as needed.
- Do not paraphrase, interpret, define, or translate any items, words, or instructions as this would be a violation of test security.
- Spell any words requested by the student.
- Adjust your reading speed and volume if requested by the student.

POST-ADMINISTRATION

- The test reader must collect scratch paper, rough drafts, and login information immediately at the end of the testing session and deliver it to the test administrator in accordance with Montana's Testing Portal policies and procedures.
- The test reader must not discuss any portion of the test with others.

ENGLISH USAGE/CONVENTIONS

- **Punctuation:** Read all text as punctuated, unless reading the text compromises the construct being measured.

- **Ellipses:** When an ellipsis is used to signify missing text in a sentence, pause briefly, and read as "dot, dot, dot."
- **Quotations:** Quotation marks should be verbalized as "quote" and "end quote" at the beginning and end of quoted material, respectively.
- **Emphasis**: When words are printed in boldface, italics, or capitals, tell the student that the words are printed that way. In order not to provide an unfair advantage to students receiving this support, test readers should be cautious not to emphasize words not already emphasized in print. Emphasis is appropriate when italics, underlining, or bold is used in the prompt, question, or answers.
- **Misspellings:** In some cases, a test item may present a word or phrase that is intentionally misspelled as part of the assessment. In these instances, the student is required to respond in a specific way. When presented with intentionally misspelled words test readers should not attempt to read the word(s) aloud as pronunciation is somewhat subjective.

IMAGES/GRAPHICS/DIAGRAMS

- Before describing an image or graphic, the test reader should determine whether the details of the picture are necessary to understanding and responding to the item(s). In many cases, an image or graphic 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. Typically, diagrams are imperative to student understanding and should be read in a logical order.
- Describe the image/graphic/diagram as concisely as possible following a logical progression. Focus on providing necessary information and ignoring the superfluous. Use grade-appropriate language when describing the image/graphic/diagram.
- Read the title or caption, if available.
- Any text that appears in the body of the image/graphic/diagram may be read to a student. Read text in images/graphics/diagrams in the order most suited for the student's needs. Often the reader moves top to bottom, left to right, in a clockwise direction, or general to specific in accordance with teaching practices.

PASSAGES

- Read the passage in its entirety as punctuated (e.g., pauses at periods and commas; raised intonation for questions). Do not verbalize punctuation marks other than ellipses and quotation marks as noted above.
- If the student requires or asks for a specific section of the passage to be re-read with the punctuation indicated, the test reader should re-read those specific lines within the passage and indicate all punctuation found within those lines as many times as requested by the student.
- When test questions refer to particular lines of a passage, read the lines referenced as though they are part of the item.

MATHEMATICAL AND SCIENTIFIC EXPRESSIONS

- The test reader must read mathematical expressions precisely and with care to avoid misrepresentation for a student who has no visual reference. For mathematics items involving algebraic expressions or other mathematical notation, it may be preferable for the reader to silently read the mathematical notations or the entire question before reading it aloud to the student.
- Test readers must read mathematical expressions with technical accuracy. Similar expressions should be treated consistently.
- In general, numbers and symbols can be read according to their common English usage for the student's grade level.
- Chemical equations, elements, and compounds should be read as printed. If an individual element is represented, the printed symbol should be read aloud; do not read the name of the element unless it is spelled out.
- Additional examples may be found in the table below.
- Abbreviations and acronyms should be read as full words. For example, 10 *cm* needs to be read as "ten centimeters." Some abbreviations may be read differently by different readers. For example, cm^3 may be read as "cubic centimeters" or "centimeters cubed.

Table 7. Test Reader Guidance for Mathematics

Numbers and Measurements		
Description	**Example(s)**	**Read as:**
Large whole numbers	632,407,981 45,000,689,112	"six hundred thirty two million, four hundred seven thousand, nine hundred eighty one" "forty five billion, six hundred eighty nine thousand, one hundred twelve"
Decimal numbers	0.056 4.37	"zero point zero five six" "four point three seven"
Fractions - common Fractions - not common - read as "numerator over denominator"	$\frac{1}{2}, \frac{1}{4}, \frac{2}{3}, \frac{4}{5}$ $\frac{14}{25}$ $\frac{487}{6972}$	"one half, one fourth, two thirds, four fifths" Other common fractions include "sixths, eighths, tenths" "fourteen over twenty five" "four hundred eighty seven over six thousand nine hundred seventy two"

Mixed numbers - read with "and" between whole number and fraction	$3\frac{1}{2}$ $57\frac{3}{4}$	"three and one-half" "fifty seven and three fourths"
Percents	62% 7.5% 0.23%	"sixty two percent" "seven point five percent" "zero point two three percent"
Money - if contains a decimal point, read as "dollars AND cents"	$4.98 $0.33 $5368.00	"four dollars and ninety eight cents" "thirty three cents" "five thousand three hundred sixty eight dollars"
Negative numbers - do NOT read negative sign as "minus"	- 3 $-\frac{5}{8}$ -7.56	"negative three" "negative five eighths" "negative seven point five six"
Dates (years)	1987 2005	"nineteen eighty seven" "two thousand five"
Roman Numerals	I II III IV	"Roman Numeral one" "Roman Numeral two" "Roman Numeral three" "Roman Numeral four"
Ratios	$x: y$	"x to y"
Square roots and cube roots	$\sqrt{6}$ $\sqrt[3]{16}$	"the square root of six" "the cube root of sixteen"
Exponents/Scientific Notation	10^{-4}	"ten to the negative fourth power"
Degrees (temperature)	32°F 0°C	"thirty-two degrees Fahrenheit" "zero degrees Celsius"
Elements	K Na	"K" "N a"
Chemical compounds	NaCl H_2O	"N a c l" "H two O"

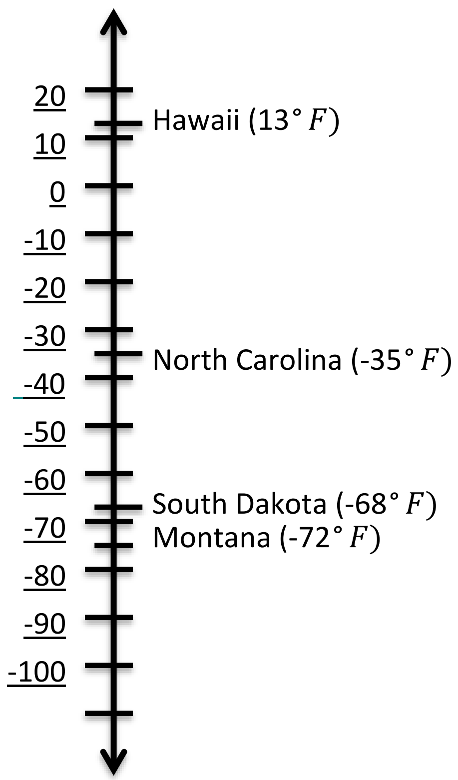
	$2CO_2$	"two C O two"
Genetics	Tt x TT	"uppercase T lowercase T, cross, uppercase T uppercase T"
Operations		
Description	**Example(s)**	**Read as:**
Addition	$\begin{array}{r}13\\ +27\\ \hline\end{array}$ 13 + 27 = 13 + 27 =?	"thirteen plus twenty seven equals" "thirteen plus twenty seven equals question mark"
Subtraction	$\begin{array}{r}487\\ -159\\ \hline\end{array}$ 487 - 159 = 487 - 159 =?	"four hundred eighty seven minus one hundred fifty nine equals" "four hundred eighty seven minus one hundred fifty nine equals question mark"
Multiplication	$\begin{array}{r}63\\ \times 49\\ \hline\end{array}$ 63 x 49 = 63 x 49 =?	"sixty three times forty nine equals" "sixty three times forty nine equals question mark"
Division – Vertical or Horizontal	$\frac{120}{15}=8$ 120 ÷ 15 = 8	"one hundred twenty divided by fifteen equals eight"
Operations with boxes	3 + □ = 8	"three plus box equals eight"
Expressions		
Description	**Example(s)**	**Read as:**
Expressions containing variables (any letter may be used as a variable)	$N + 4$ 8x — 3 4 (y — 2) + 5 =7 $V = \frac{4}{3}\pi r^3$	"'N' plus four" "eight 'x' minus three" "four open parenthesis 'y' minus two close parenthesis plus five equals seven"

	$\frac{\lvert t\rvert - 2}{6} \le 15$ $x^2y^3 = -36$ $156x \ge 4$	"'V' equals four thirds pi 'r' cubed" "the absolute value of 't' (pause) minus two (pause) over six is less than or equal to fifteen" "'x' squared 'y' cubed equals negative thirty six" or "'x' to the second power times 'y' to the third power equals negative thirty six" "one hundred fifty six 'x' is greater than or equal to four"
Functions and inverse functions (Read "of" instead of parentheses)	$f(x)$ $f(x+2)$ $f(g(x))$	"F of x" "F of x plus 2" "F of g of x"
Coordinate pairs Answer choices with no other text	the point (-1, 2) the point A is at (6, 3) A. (-3, -4)	"the point (pause) negative one comma two" "the point 'A' is at (pause) six comma three" "'A' (pause) negative three comma negative four"
Chemical reactions and equations	2 HCl + 2 Na -> 2 NaCl + H_2	"two 'H C l' plus two 'N a' yields 'N a C l' plus H two"
Comparing Lines, Shapes, and Angles		
Description	**Example(s)**	**Read as:**
Parallels	$\overline{AB} \mid\mid \overline{CD}$	"line segment AB is parallel to line segment CD"
Perpendiculars	$\overline{AB} \perp \overline{CD}$	"line segment AB is perpendicular to line segment CD"
Similar and congruent	$\Delta ABC \sim \Delta DEF$ $\angle ABC \cong \angle DEF$	"triangle A B C is similar to triangle D E F" "angle A B C is congruent to angle D E F"

Lines, line segments, rays, arcs	$\overleftrightarrow{BC}$ $\overline{CD}$ $\overrightarrow{BC}$ $\overset{\frown}{BC}$	"line B C" "line segment C D" "ray B C" "arc B C"
Trigonometry		
Description	**Example(s)**	**Read as:**
Sine	sin25°	"sine twenty five degrees"
Cosine	cos35°	"cosine thirty five degrees"
Tangent	tan10°	"tangent ten degrees"

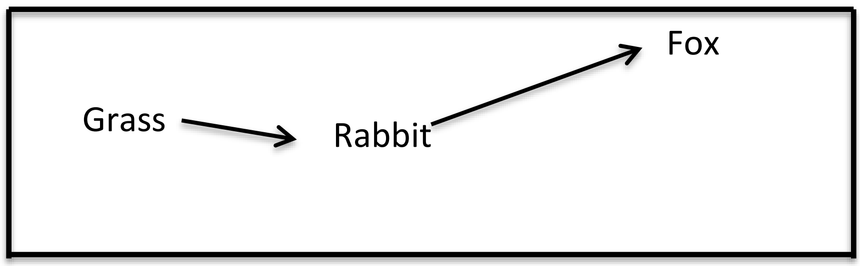
Images/Graphics/Diagrams/Tables

From top to bottom



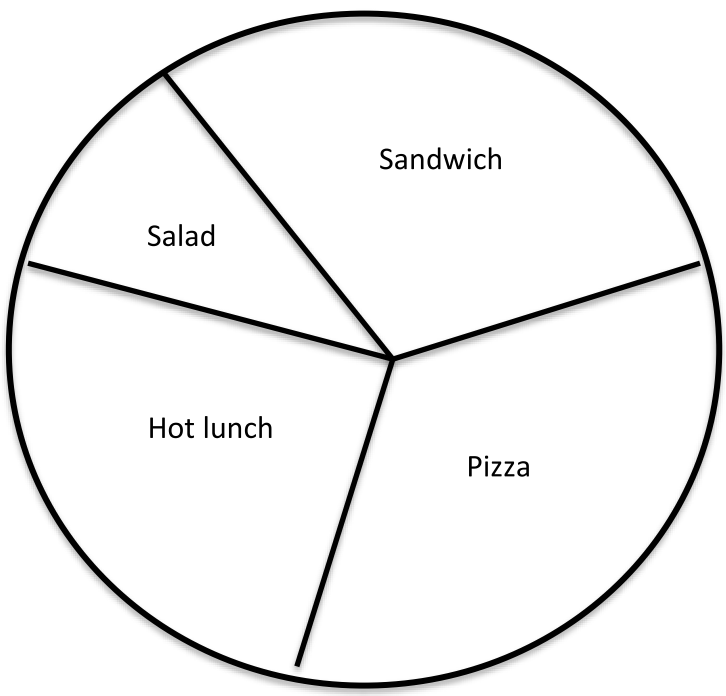
"From top to bottom the figure is labeled: Hawaii thirteen degrees Fahrenheit, North Carolina negative thirty five degrees Fahrenheit, South Dakota negative sixty eight degrees Fahrenheit, Montana negative seventy two degrees Fahrenheit"

FROM LEFT TO RIGHT



"From left to right, the figure reads: Grass, Rabbit, Fox"

CLOCKWISE (START WHEREVER MAKES SENSE.)



"Clockwise from the top, the figure reads: Sandwich, Pizza, Hot lunch, Salad"

TABLES

1. Read title.
2. Total up the columns and rows.

3. Read column/row headings
4. Read cell values (only as directional language for the first one)

Results from School Walk-a-Thon

Number of Students	*Number of miles walked*
30	112
46	214
37	98
41	189

"The title of the table is Results from School Walk-a-Thon. The table has 2 columns and 4 rows. From left to right, the column headings read Number of Students, Number of Miles Walked. From left to right the first row reads thirty, one hundred twelve. The second row reads forty six, two hundred fourteen. The third row reads thirty seven, ninety eight. The fourth row reads forty one, one hundred eighty nine.

SUGGESTED TEST READER SCRIPT (TO BE USED WITH STUDENT IN ADVANCE OF THE DAY OF TESTING)

Hi,
I'm the person who will be reading your test to you when you take your assessment next week in [science]. I wanted to let you know how we'll work together. When I'm reading a test to you, it's very different from when I'm reading to you during class time. I have to follow certain rules.

- I cannot help you with any answers.
- I cannot click on anything on the screen.[1]
- I will not be using different character voices or changes in my tone when I read. I will be using a very direct voice that does not change very much, no matter how exciting the story or test item gets.
- If there is a picture that has words in it, I will read those words. If you ask, I will re-read the words as well.
- Sometimes there may be something about a word or phrase that might give you a hint if I read it out loud. In those cases, I will skip the word, point to it on screen [**or on your booklet if braille or print on request], and continue to read.
- I can still help you with your [**list any assistive technology that the student may require that would need adult support -- if that support is provided by you].
- You can ask me to re-read parts of the test if you didn't hear me or need more time to think.
- You can ask me to pause my reading if you need to take a break.

[1] A reader may click on something on the screen only if this is an identified need in the student's IEP or 504 plan and the reader has received appropriate training on when and how to do so.

- You can ask me to slow down or speed up my reading or read louder or softer if you are having trouble understanding what I read.
- I will only read certain types of punctuation, but if you need me to re-read a sentence and tell you how it was punctuated, I can do that.
- If you ask me a question about the test all I will say is: "Do your best work. I cannot help you with that."
- Do you have any questions for me about how we’ll work together during the test?

APPENDIX C. SCRIBING GUIDELINES

A scribe is an adult who writes down what a student dictates in a variety of ways (e.g., speech, American Sign Language (ASL), braille, assistive communication device). The guiding principle in scribing is to ensure that the student has access to and is able to respond to test content. Scribes are allowable on the Montana Science Assessment as a designated support. For information on documentation requirements and decision-making criteria for use of scribes and all other supports please see the *Usability, Accessibility, and Accommodations Guidelines.*

QUALIFICATIONS FOR SCRIBES

- The scribe should be an adult who is familiar with the student, such as the teacher or teaching assistant who is typically responsible for scribing during educational instruction and assessments.
- Scribes must have demonstrated knowledge and experience in the subject for which scribing will be provided.
- Scribes should have extensive practice and training in accordance with security policies and procedures as articulated in Montana Office of Public Instruction test administration manuals, guidelines, and related documentation.

PREPARATION

- Scribes should read and sign a test security/confidentiality agreement prior to test administration. If scribes have a TIDE account and have taken the TA Certification Course, then their Test Security Agreement (TSA) will have been signed online. If not, the System Test Coordinator (STC) should ensure that the test reader signs a paper copy of the TSA.
- Scribes are expected to familiarize themselves with the test format in advance of the scribing session. Having a working familiarity with the test environment will help facilitate the scribe's ability to record the student's answers. Scribes may wish to review the practice test to become familiar with the assessment.
- Scribes should be familiar with the Individualized Education Program (IEP) or 504 plan if the student for whom they are scribing has a disability, so that there are plans in place for providing all needed designated supports and accommodations.
- Scribes should also have a strong working knowledge of the embedded and non-embedded universal tools, designated supports, and accommodations available on assessments.
- Scribes should review the *Scribing Protocol* with the student at least one to two days prior to the test event.
- Scribes should practice the scribing process with the student at least once prior to the scribing session.

GENERAL GUIDELINES

- Scribing must be administered so that the interaction between a scribe and a student does not interrupt other test takers, or inadvertently reveal the student's answers.
    - If not in a separate setting, the scribe should be situated near enough (adhering to local health and safety protocol) to the student to prevent their conversations from reaching other students in the room.
- For computer-based administrations, scribes must enter student responses directly into the test interface, making use of the embedded and non-embedded tools available for a given item and student.
- Scribes are expected to comply with student requests regarding use of all available features within the test environment.
- Scribes may respond to procedural questions asked by the student (e.g., test directions, navigation within the test environment, etc.).
- Scribes may not respond to student questions about test items if their responses compromise validity of the test. The student must not be prompted, reminded, or otherwise assisted in formulating his or her response during or after the dictation to the scribe.
- Scribes may ask the student to restate words or parts as needed. Such requests must not be communicated in a manner suggesting that the student should make a change or correction.
- Scribes may not question or correct student choices, alert students to errors or mistakes, prompt or influence students in any way that might compromise the integrity of student responses. A scribe may not edit or alter student work in any way and must record exactly what the student has dictated.
- Students must be allowed to review and edit what the scribe has written. If necessary, the student can request the scribe to read aloud the completed text before final approval.

CONSIDERATIONS FOR STUDENTS ALSO USING ASL OR OTHER SIGN SYSTEM

- The scribe should be proficient in the sign system utilized (e.g., ASL) or the scribe should be working with an interpreter proficient in the sign system, as determined by OPI.
- When a constructed response is required, the interpreter/scribe should convey the meaning behind the student's indicated response.
- The interpreter/scribe should show the student the written response, but NOT sign the response to the student.
    - Probing or clarifying is allowed in the case of classifiers for students using ASL or other sign systems.
- Students may review the written or typed response on paper or on the computer screen and indicate any changes or revisions to the scribe.

CONSIDERATIONS FOR STUDENTS USING BRAILLE

- The scribe should be proficient in reading (visually or tactually) braille in all braille codes used by the student, as determined by OPI.
- The scribe should enter the responses on paper or online exactly as the student has brailled. In addition to following the content-specific guidelines above, errors in braille code should not be corrected.
- The scribe may ask for the student to read back brailled responses for clarification if the brailled response is difficult to read due to student corrections.
- Students may review the written or typed response on paper or on the computer screen by either using the scribe to read back the entered response or using assistive technology. Students may indicate any changes or revisions to the scribe.

POST-ADMINISTRATION

- The scribe will submit online or paper-based student responses and collect scratch paper, rough drafts, and login information immediately at the end of the testing session and deliver it to the test administrator in accordance with Montana Office of Public Instruction policies and procedures.

Montana Science Assessment

2022–2023

Volume 6:
Score Interpretation Guide

OFFICE OF PUBLIC INSTRUCTION
ELSIE ARNTZEN, STATE SUPERINTENDENT

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

1. MONTANA SCORE REPORTS

In spring 2023, the Montana Science Assessment was administered to Montana students in grades 5 and 8. The purpose of the *Score Interpretation Guide* (Volume 6 of the *Montana Science Assessment 2022–2023 Technical Report*) is to document the features of the Montana Reporting System, which is designed to assist stakeholders in reviewing and downloading test results and understanding and appropriately using these results. Additionally, this volume describes the score types reported for the spring 2023 assessments, the appropriate uses of those score types, the inferences that can be drawn from them, and features of the score report.

1.1 OVERVIEW OF THE MONTANA SCORE REPORTS

The Montana Science Assessment was first administered operationally statewide in spring 2022. Test scores from the spring 2022 assessment were provided to districts and schools through the Montana Reporting System on September 12, 2022, after the standard-setting workshop that occurred on August 2–August 3, 2022. The Reporting System provided information on student performance and aggregated summaries at several levels—district, school, and roster.

The Montana Reporting System (https://mt.reporting.cambiumast.com) is a web-based application that provides Montana Science Assessment results at various levels. Test results are available for users based on their roles and the privileges they receive based on the authentication granted to them. There are four basic levels of user roles, including the state, district, school, and teacher levels. Each user is granted drill-down access to system reports based on his or her assigned role. This means that teachers can access data for their roster(s) of students only, schools can access data for the students in their school only, and districts can access data for all schools and students in their district only.

The following users have access to the system:

- **State Users.** Access to all data at the state, district, school, teacher, and student levels
- **Authorized Representative (AR) and System Test Coordinator (STC) Users.** Access to all data for their district and the schools and students in their district
- **Building Coordinator (BC) Users.** Access to all data for their school and the students in their school
- **Test Administrator (TA) Users.** Access to all aggregate data for their rosters and the students within their rosters

Access to reports is password protected, and users can access data at and below their assigned level. For example, an BC user can access the school report of students for his or her school but not for another school.

1.2 OVERALL SCORES AND DISCIPLINE-LEVEL SCORES

Each student receives a single scale score for each subject tested if there is a valid score to report. A student's score is based only on the operational items on the assessment. A scale score is used to describe how well a student performed on a test and can be interpreted as an estimate of the student's knowledge and skills measured. The scale score is transformed from a theta score, which is estimated based on mathematical models. Low scale scores can be interpreted as an indication that the student does not possess sufficient knowledge and skills measured by the test. Conversely, high scale scores can be interpreted as an indication that the student has proficient knowledge and skills measured by the test. Interpretation of scale scores is more meaningful when the scale scores are used along with performance levels and Performance-Level Descriptors (PLDs).

Based on the scale score, students will receive an overall performance level. Performance levels are proficiency categories on a test which students fall into based on their scale scores. For the Montana Science Assessment, scale scores are mapped into four performance levels:

1. Level 1
2. Level 2
3. Level 3
4. Level 4

For details on the standard-setting process, refer to Volume 3, Setting Performance Standards, of this technical report.

PLDs are a description of content area, knowledge, and skills that students at each performance level are expected to possess. Thus, performance levels can be interpreted based on PLDs. Generally, students performing on the Montana Science Assessment at Levels 3 and 4 are considered on track to demonstrating progress toward mastery of the knowledge and skills necessary for college and career readiness.

In addition to an overall score, students receive discipline-level scores. For the Montana Science Assessment, student performance on each discipline level is reported at three performance categories:

1. Below Standard
2. At/Near Standard
3. Above Standard

Unlike the performance levels for the overall test, student performance on each of the discipline levels is evaluated with respect to the Level 3 performance standard. Student performance at either Below Standard or Above Standard can be interpreted as student performance clearly below or above the Level 3 cut score for a specific discipline. Student performance at At/Near Standard can be interpreted as student performance that does not provide enough information to tell whether students reached the Level 3 mark for the specific discipline. Volume 3 of this technical report provides detail on the standard setting procedures.

Table 1 displays the disciplines and discipline-level claims for science, by grade.

Table 1. Disciplines and Discipline-Level Claims for Science

Grade	Discipline	Claim
5, 8	Practices and Concepts in Life Sciences	The student is able to use science and engineering practices to demonstrate understanding of the disciplinary core ideas and crosscutting concepts in Life Science.
	Practices and Concepts in Physical Sciences	The student is able to use science and engineering practices to demonstrate understanding of the disciplinary core ideas and crosscutting concepts in Physical Science.
	Practices and Concepts in Earth and Space Sciences	The student is able to use science and engineering practices to demonstrate understanding of the disciplinary core ideas and crosscutting concepts in Earth and Space Science.

1.3 Montana Reporting System

The Montana Reporting System generates a set of online score reports that describes student performance for students, families, educators, and other stakeholders. The online score reports are produced after the tests are submitted by the students, handscored and machine-scored, and finally, processed into the Reporting System. In addition to each individual student's score report, the Reporting System produces aggregate score reports for teachers, schools, districts, and states.

Furthermore, to facilitate comparisons, each aggregate report contains the summary results for the selected aggregate unit, as well as all aggregate units above the selected aggregate. For example, if a school is selected, the summary results of the district to which the school belongs are provided so that the school performance can be compared with the district performance. If a teacher is selected, the summary results for the school and the district above the teacher are also provided for comparison purposes.

1.4 Available Reports on the Montana Reporting System

The Montana Reporting System is hierarchically structured. An authorized user is able to view reports at his or her own aggregated unit and any lower level of aggregation. For example, a school user can view only the reports and data at the school and student levels of his or her school. AR users can view the reports and data for their districts and the student-level results for all of their schools.

Table 2 summarizes the types of score reports that are available in the Reporting System and the levels at which the reports can be viewed. A description of each report is also provided. Data files are also accessible for district to download.

For detailed information on available reports and available features, educators can refer to the *Reporting System User Guide*, included in Appendix 6-A, Reporting System User Guide.

Table 2. Montana Score Reports Summary

Report	Description	Unit of Aggregation				
		State	District	School	Roster	Student
Summary Performance	Summary of performance (to date) across grades and subjects or courses for the current administration	✓	✓	✓	✓	
Aggregate-Level Subject Report	Summary of overall performance for a subject and grade for all students in the defined level of aggregation	✓	✓	✓	✓	
Aggregate-Level Discipline-Level Score Report	Summary of overall performance on each discipline level for each grade across all students within the selected level of aggregation	✓	✓	✓	✓	
Aggregate-Level Disciplinary Core Ideas Report	Summary of overall performance on each disciplinary core idea for a given subject and grade across all students within the selected level of aggregation	✓	✓	✓	✓	
Student-Level Subject Report	List of all students who belong to a school, teacher, or roster with their associated subject or course scores for the current administration			✓	✓	✓
Student-Level Discipline-Level Score Report	List of all students who belong to a school, teacher, or roster with their associated discipline-level performance for the current administration			✓	✓	✓
Individual Student Report (ISR)	Detailed information about a selected student's performance in a specified subject or course, including overall subject and discipline-level results					✓
Data Files	Text/CSV file containing overall and discipline-level scale scores and performance levels along with demographic information		✓	✓	✓	✓

1.4.1 Reporting by Subgroup

The aggregate score reports provide overall student results by default but can at any time be analyzed by subgroups based on demographic data. When used on aggregate-level reports, an additional level of analysis will be provided by aggregating students based on subgroup. For example, when the Gender subgroup is selected, the Reporting System will display aggregate results for all students, male and female. When used on student-level reports, subgroups can instead be used to filter individual results. For example, a user will have the option to select Male or Female after the Gender subgroup is selected.

Users can see student assessment results by any subgroup at any time by selecting the desired subgroup from the *Breakdown By* drop-down menu available. Table 3 presents the types of subgroups and subgroup categories provided in the Reporting System. The Reporting System allows users to select up to three subgroups for comparison.

Table 3. Montana List of Subgroups

Breakdown by Category	Displayed Category
Race/Ethnicity	Two or More Races
	American Indian or Alaskan Native
	Asian
	Hispanic or Latino
	Black or African American
	White
	Pacific Islander
Gender	Male
	Female
IDEA (Individuals with Disabilities Education Act) Indicator	Yes
	No
English Learner Status	Yes
	No
Section 504 Plan Status	Yes
	No
Economic Disadvantage Status	Yes
	No
Enrolled Grade	Grade 5
	Grade 8

1.4.2 Summary Performance Report

Homepage-authorized users can log on to the Reporting System dashboard to view summaries of students' performance across grades and subjects. Using the dashboard, state personnel and district personnel can access district summaries, school personnel can access school summaries, and teachers can access summaries of their students. The Summary Performance Distribution has the following features:

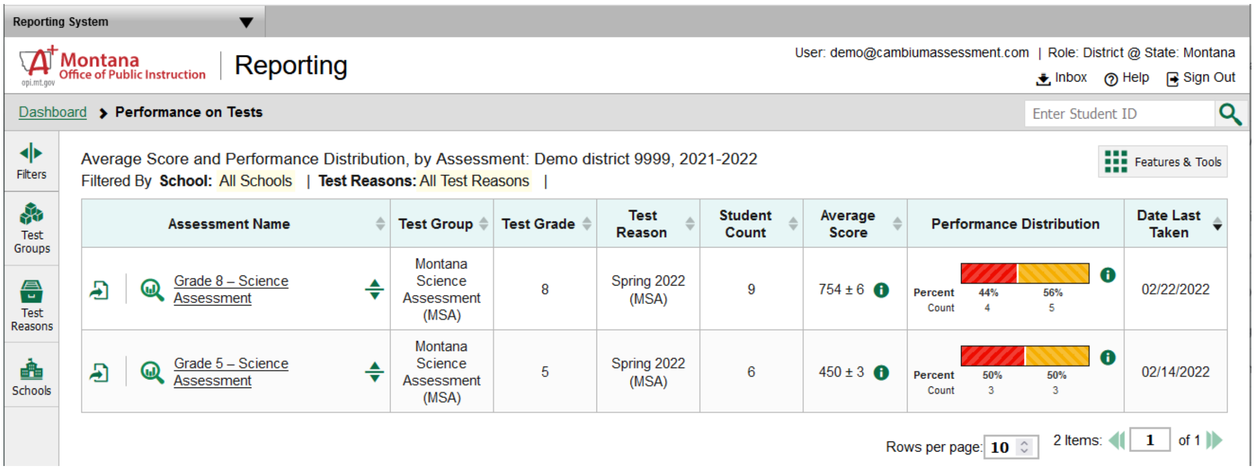
- Displays summary data separated by grade and subject
- Presents level of aggregation based on a user's role
- Reports number of students tested, average scale score and standard error of the average scale score, and percentage of students in each performance level 1-4.

The four performance levels are color-coded in the Performance Distribution bar as follows:

1. Red is the percentage of "Level 1" students.
2. Orange is the percentage of "Level 2" students.
3. Green is the percentage of "Level 3" students.
4. Blue is the percentage of "Level 4" students.

Figure 1 presents sample Summary Performance Reports at the district level.

Figure 1. District-Level Summary Performance Report



1.4.3 Aggregate-Level Subject Report

Detailed summaries of student performance within a grade and subject area are made available in the Aggregate-Level Subject Report. The Aggregate-Level Subject Report presents results for the aggregate unit, as well as the results for any higher-level aggregate units. For example, a school's Aggregate-Level Subject Report will contain the summary results of the school's district so that school performance can be compared with district performance.

The Aggregate-Level Subject Report provides the aggregate summaries on a specific subject area, including

- number of students;
- average scale score and standard error of the average scale score;
- percentage of students in each performance level 1-4 (the "Performance Distribution column"); and
- percentage of students meeting standard across levels 3 and 4 (the "Percent Perficient" column).

The summaries are also presented for overall students and by subgroups. Figure 2 presents an example of Aggregate-Level Subject Reports for grade 8 science at the district level without subgroups. Figure 3 highlights grade 8 science at the district level when a user selects a subgroup of gender.

Figure 2. District Aggregate-Level Subject Report for Grade 8 Science

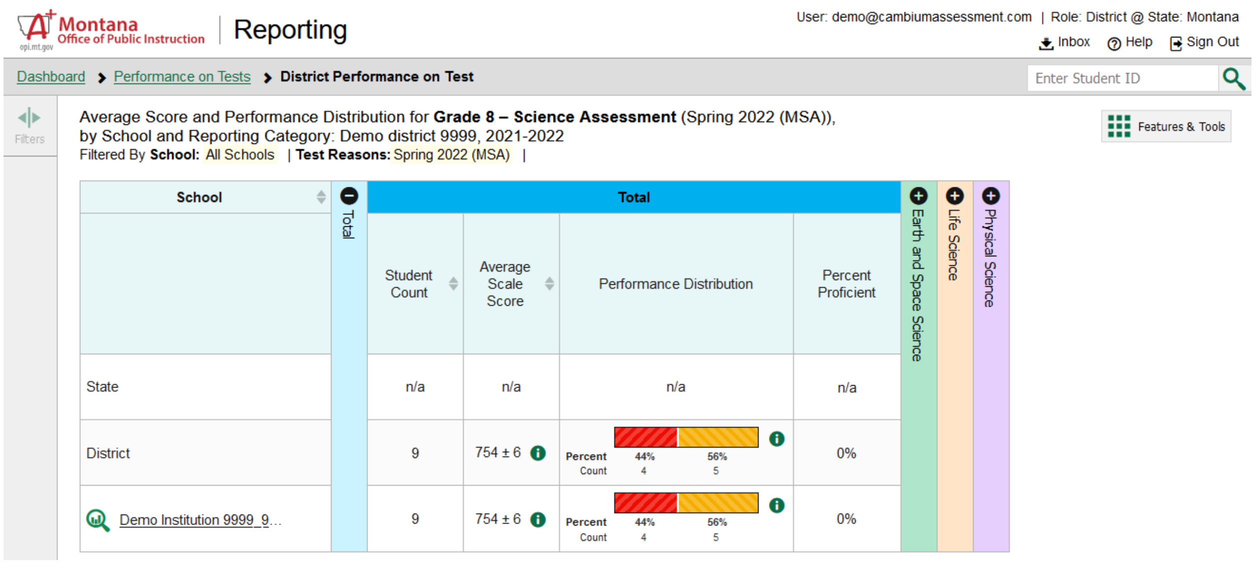
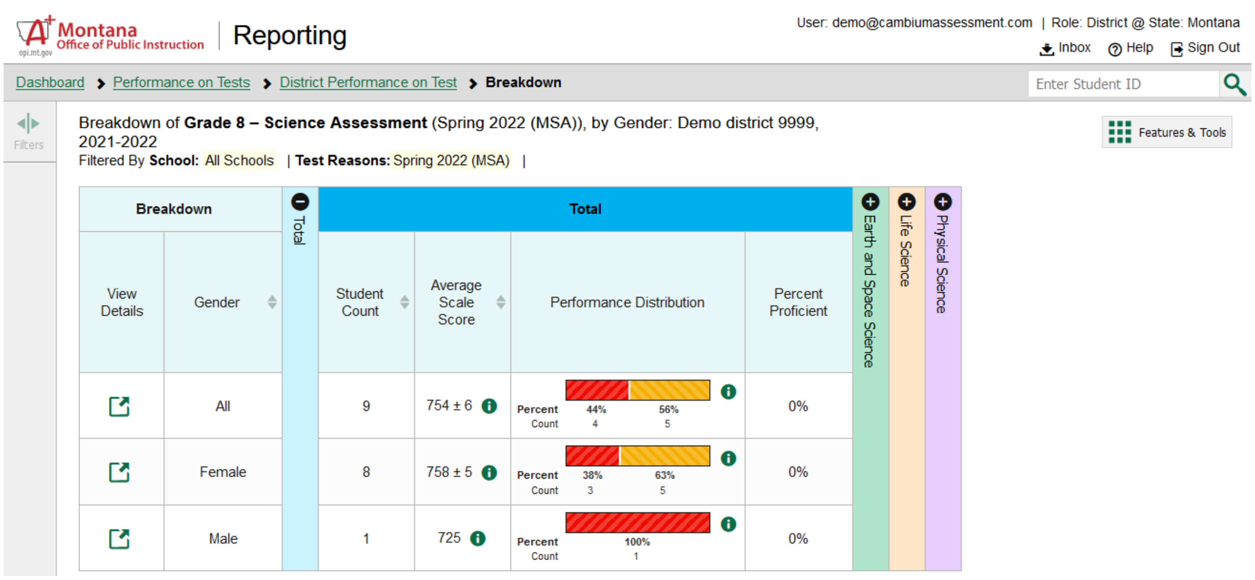


Figure 3. District Aggregate-Level Subject Report for Grade 8 Science by Gender



1.4.4 Aggregate-Level Discipline-Level Report

The Aggregate-Level Discipline-Level Report provides the aggregate summaries on student performance in each discipline level for each grade. The summaries on the Aggregate-Level Discipline-Level Report include

- number of students;
- average discipline scale score and standard error of the average discipline scale score;
- percentage of students in each performance category for each of the disciplines.

The three discipline-level performance levels are color-coded in the performance distribution bar as follows:

1. Red is the percentage of "Below Standard" students.
2. Orange is the percentage of "At/Near Standard" students.
3. Green is the percentage of "Above Standard" students.

The Aggregate-Level Disciplinary Core Ideas-Level Report lists data on the performance of student groups on each standard of a subject for the current window and reports Areas Where Performance Indicates Proficiency and Areas of Strongest and Weakest Performance. For Areas Where Performance Indicates Proficiency, the Performance Distribution column produces information on how a group of students in a class, school, or district performed on the standard compared to the proficiency cuts. It shows whether performance on this standard for this group was above, no different than, or below what is expected of students at the proficient level. This indicator shows strengths and weaknesses for a group of students and is provided only at an aggregate level since it is unstable at the individual level.

For Areas of Strongest and Weakest Performance, the expected performance is determined based on the students' overall performance on the entire test. The key for the "Proficient?" and "Weak or Strong?" symbols is provided below.

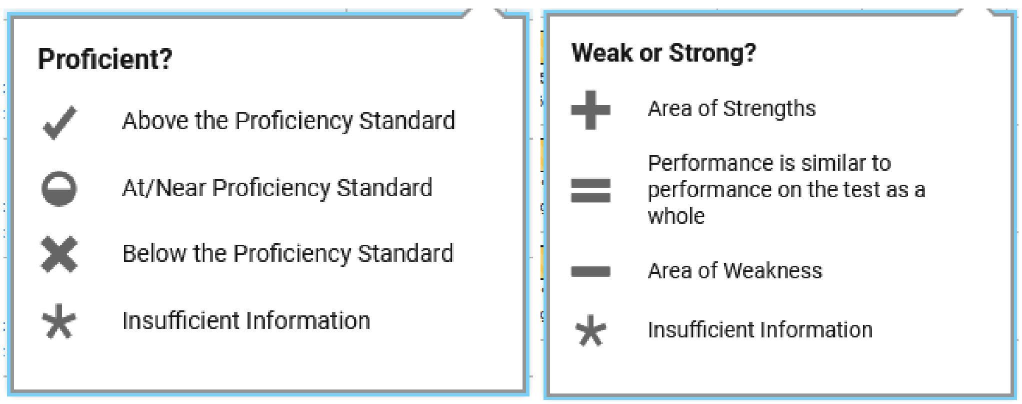
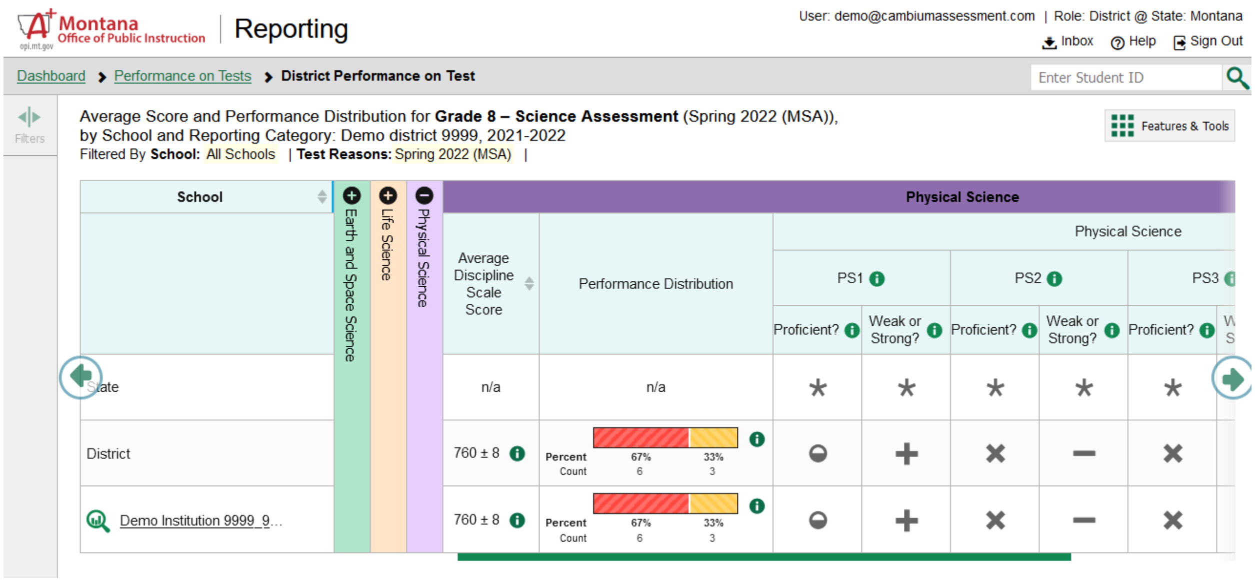


Figure 3 presents an example of the District Aggregate-Level Discipline-Level Report for grade 8 science.

Figure 3. District Aggregate-Level Discipline-Level Report for Grade 8 Science



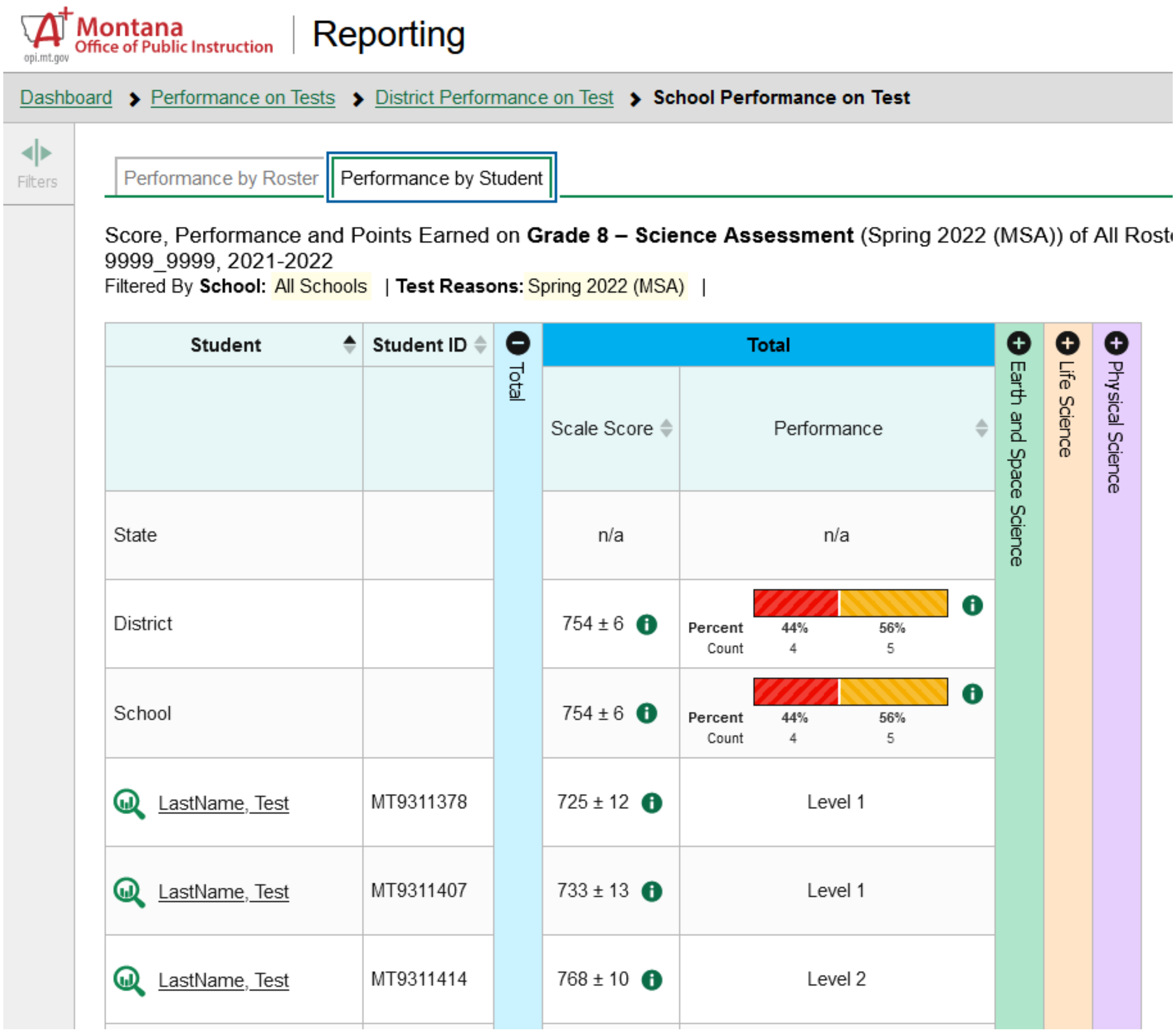
1.4.5 Student-Level Subject Report

The Student-Level Subject Report lists all students who belong to the selected aggregate level, such as a school, and reports the following measures for each student:

- Scale score
- Overall subject performance level

Figure 4 demonstrates an example of the Student-Level Subject Report for grade 8 science.

Figure 4. Student Roster Subject Report for Grade 8 Science



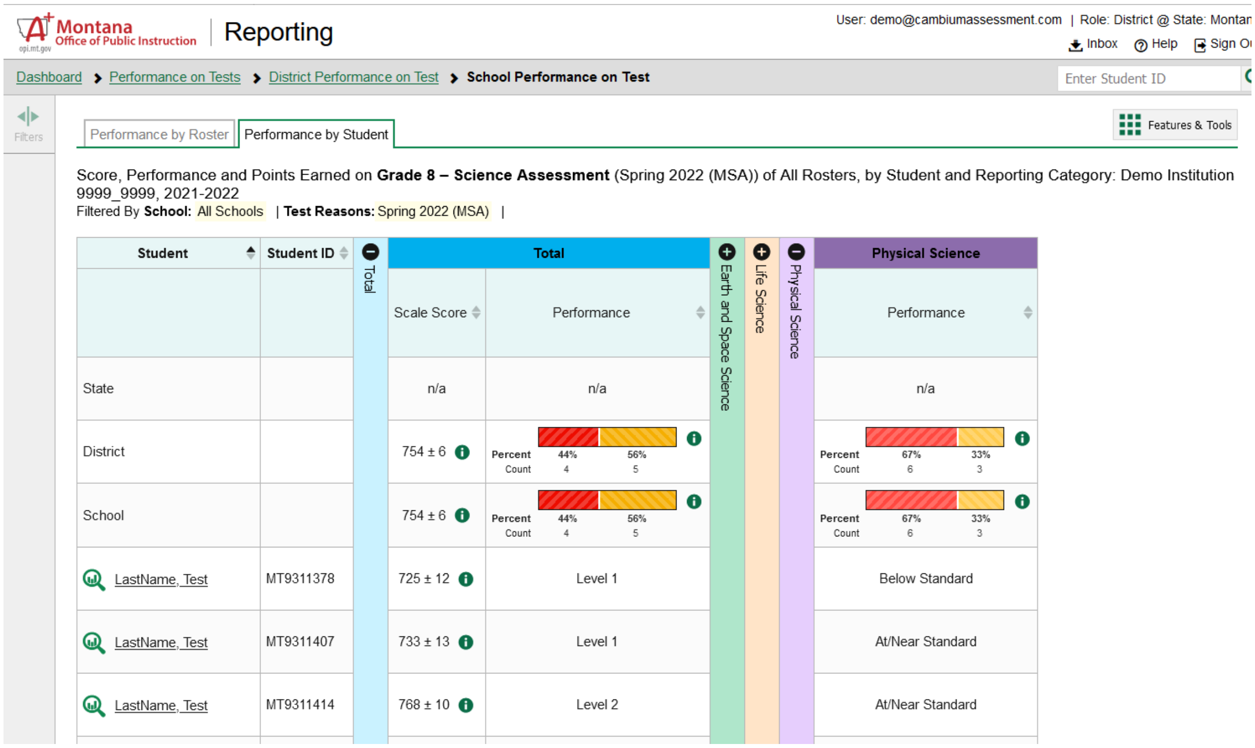
1.4.6 Student-Level Discipline-Level Report

The Student-Level Discipline-Level Report lists all students who belong to the selected aggregate level, such as a school, and reports the following measures for each student:

- Scale score
- Overall subject performance level
- Discipline performance category (i.e., Earth and Space Sciences, Life Sciences, and Physical Sciences)

Figure 5 presents an example of the Student-Level Discipline Report for grade 8 science.

Figure 5. Student Roster Discipline Report for Grade 8 Science



1.4.7 Individual Student Report

When a student receives a valid test score, an Individual Student Report (ISR) can be generated in the Reporting System. The ISR contains the following measures:

- Scale score and standard error of measurement (SEM)
- Overall subject performance level
- Average scale scores for student's district and school
- Performance category in each science discipline

At the top of the report, information includes:

- Student's name
- Scale score with SEM
- Performance level

In the middle section of the report, information includes:

- Barrel chart with student's scale score and SEM (using a sign of "±")
- PLDs with cut scores at each performance level
- Average scale scores and standard errors for district and school aggregation levels

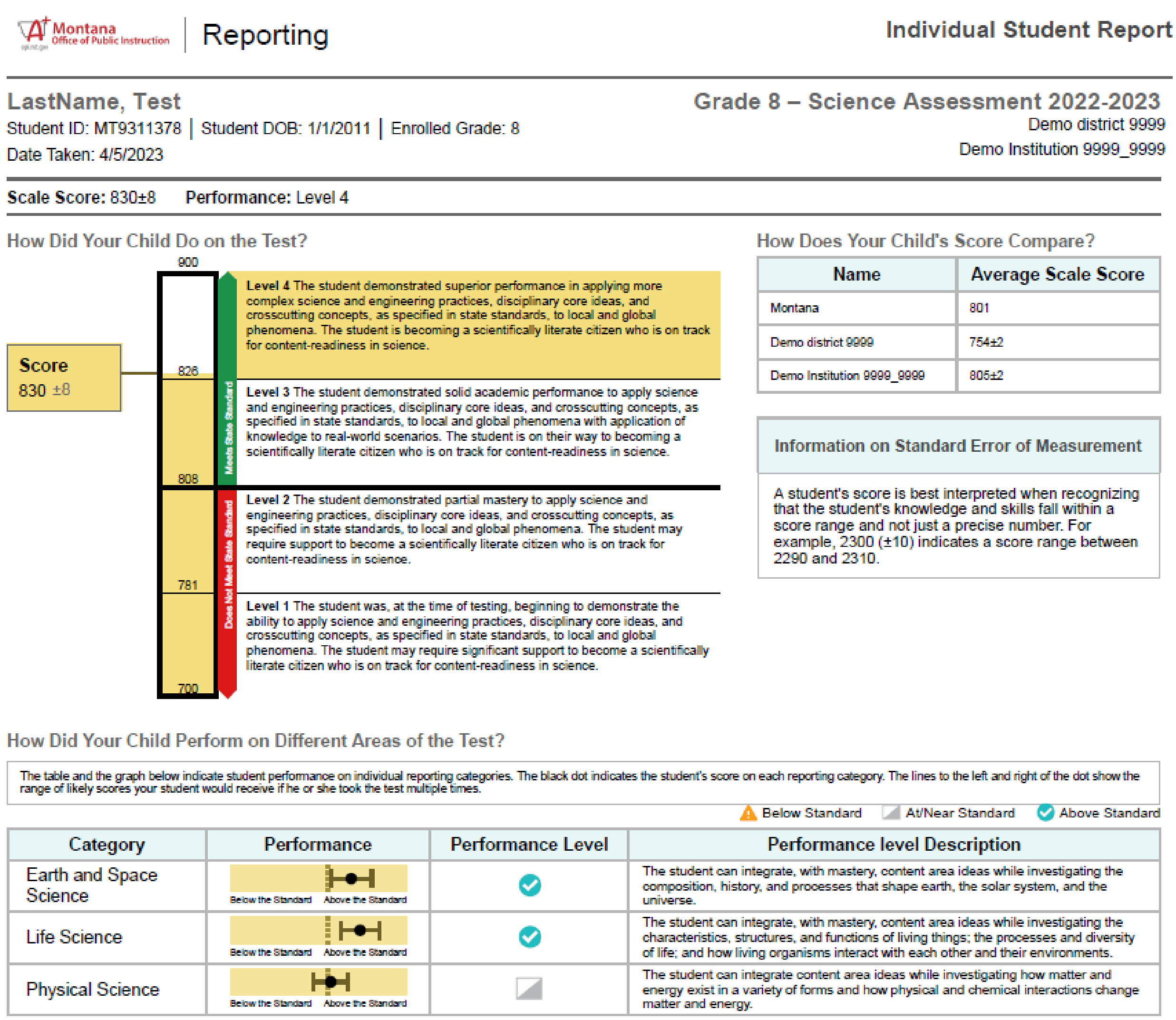
- Note: the "±" next to the student's scale score is the SEM of the scale score, whereas the "±" next to the average scale scores for aggregate levels represents the SEM of the average scale scores.

At the bottom of the report, information includes:

- Detailed information on student performance on each discipline level
  - Note: Bar charts in the Discipline table below the overall scale score barrel chart ("How Did Your Child Perform on Different Areas of the Test?") show how students performed on each discipline (the black line/dot), relative to the discipline-level performance standard (dashed white line within the yellow bar).

Figure 6 presents an example of the ISR for grade 8 science.

Figure 6. Individual Student Report for Grade 8 Science



Montana Office of Public Instruction | Reporting

Individual Student Report

LastName, Test

Student ID: MT9311378 | Student DOB: 1/1/2011 | Enrolled Grade: 8

Date Taken: 4/5/2023

Grade 8 – Science Assessment 2022-2023

Demo district 9999

Demo Institution 9999_9999

Scale Score: 830±8 **Performance:** Level 4

How Did Your Child Do on the Test?

Score 830 ±8

900, 826, 808, 781, 700

Meets State Standard / Does Not Meet State Standard

Level 4 The student demonstrated superior performance in applying more complex science and engineering practices, disciplinary core ideas, and crosscutting concepts, as specified in state standards, to local and global phenomena. The student is becoming a scientifically literate citizen who is on track for content-readiness in science.

Level 3 The student demonstrated solid academic performance to apply science and engineering practices, disciplinary core ideas, and crosscutting concepts, as specified in state standards, to local and global phenomena with application of knowledge to real-world scenarios. The student is on their way to becoming a scientifically literate citizen who is on track for content-readiness in science.

Level 2 The student demonstrated partial mastery to apply science and engineering practices, disciplinary core ideas, and crosscutting concepts, as specified in state standards, to local and global phenomena. The student may require support to become a scientifically literate citizen who is on track for content-readiness in science.

Level 1 The student was, at the time of testing, beginning to demonstrate the ability to apply science and engineering practices, disciplinary core ideas, and crosscutting concepts, as specified in state standards, to local and global phenomena. The student may require significant support to become a scientifically literate citizen who is on track for content-readiness in science.

How Does Your Child's Score Compare?

Name	Average Scale Score
Montana	801
Demo district 9999	754±2
Demo Institution 9999_9999	805±2

Information on Standard Error of Measurement

A student's score is best interpreted when recognizing that the student's knowledge and skills fall within a score range and not just a precise number. For example, 2300 (±10) indicates a score range between 2290 and 2310.

How Did Your Child Perform on Different Areas of the Test?

The table and the graph below indicate student performance on individual reporting categories. The black dot indicates the student's score on each reporting category. The lines to the left and right of the dot show the range of likely scores your student would receive if he or she took the test multiple times.

Below Standard | At/Near Standard | Above Standard

Category	Performance	Performance Level	Performance level Description
Earth and Space Science	Below the Standard / Above the Standard	Above Standard	The student can integrate, with mastery, content area ideas while investigating the composition, history, and processes that shape earth, the solar system, and the universe.
Life Science	Below the Standard / Above the Standard	Above Standard	The student can integrate, with mastery, content area ideas while investigating the characteristics, structures, and functions of living things; the processes and diversity of life; and how living organisms interact with each other and their environments.
Physical Science	Below the Standard / Above the Standard	At/Near Standard	The student can integrate content area ideas while investigating how matter and energy exist in a variety of forms and how physical and chemical interactions change matter and energy.

1.4.8 Data File

Reporting System users have the option to quickly generate a comprehensive data file of their students' scores. Data files (See Figure 7) can be downloaded in Microsoft Excel or CSV format and contain a wide variety of data, which includes scale scores, reporting discipline scores, demographic data, and performance levels. Data files can be useful as a resource for further analysis. Data files can be generated at the district, school, teacher, or roster level. Here is a list of the data variables provided in the data file.

- Section 504 Status
- Enrolled Grade
- Race/Ethnicity
- Gender
- Limited English Proficiency Status
- IDEA Indicator
- Economic Disadvantage Status
- Enrolled District
- Enrolled School
- Test Reason
- Test OppNumber
- Date Taken
- Test Completion Date
- Grade 8 – Science Assessment Scale Score
- Grade 8 – Science Assessment Scale Score Standard Error
- Grade 8 – Science Assessment Performance
- Earth and Space Science Performance
- Life Science Performance
- Physical Science Performance

Figure 7. Data File

D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X
Section 50	Enrolled G	Race/Ethn	Gender	Limited Er	IDEA Indic	Economic	Enrolled D	Enrolled S	Test Reas	Test Opp	Date Take	Test Comp	Grade 8 –	Grade 8 –	Grade 8 –	Earth and	Life Scien	Physical Science Performance		
No	9	White	Male	No	No	No	Demo dist	Demo Inst	Spring 202	Most Rece	04/05/202	04/05/202	802	8	Level 2	At/Near S	At/Near S	At/Near Standard		
No	9	White	Female	No	No	No	Demo dist	Demo Inst	Spring 202	Most Rece	05/08/202	05/09/202	791	8	Level 2	At/Near S	Below Sta	At/Near Standard		
No	9	White	Female	No	No	No	Demo dist	Demo Inst	Spring 202	Most Rece	04/05/202	04/05/202	830	8	Level 4	Above Sta	Above Sta	At/Near Standard		
No	9	White	Male	No	Yes	Yes	Demo dist	Demo Inst	Spring 202	Most Rece	04/04/202	04/06/202	773	9	Level 1	Below Sta	Below Sta	Below Standard		
No	9	White	Male	No	No	No	Demo dist	Demo Inst	Spring 202	Most Rece	04/05/202	04/05/202	791	9	Level 2	At/Near S	At/Near S	Below Standard		
No	9	White	Male	No	No	No	Demo dist	Demo Inst	Spring 202	Most Rece	04/05/202	04/05/202	803	8	Level 2	At/Near S	At/Near S	At/Near Standard		
No	9	Multi-Raci	Male	No	Yes	No	Demo dist	Demo Inst	Spring 202	Most Rece	04/06/202	04/06/202	789	9	Level 2	At/Near S	At/Near S	Below Standard		
No	9	White	Female	No	No	No	Demo dist	Demo Inst	Spring 202	Most Rece	04/05/202	04/05/202	798	9	Level 2	At/Near S	Below Sta	At/Near Standard		
Yes	9	White	Male	No	No	No	Demo dist	Demo Inst	Spring 202	Most Rece	04/04/202	04/04/202	757	9	Level 1	Below Sta	Below Sta	Below Standard		
No	9	White	Female	No	No	No	Demo dist	Demo Inst	Spring 202	Most Rece	04/04/202	04/06/202	824	8	Level 3	At/Near S	At/Near S	Above Standard		
No	9	White	Female	No	No	No	Demo dist	Demo Inst	Spring 202	Most Rece	04/04/202	04/04/202	824	10	Level 3	At/Near S	At/Near S	At/Near Standard		
No	9	White	Male	No	Yes	Yes	Demo dist	Demo Inst	Spring 202	Most Rece	04/04/202	04/04/202	778	9	Level 1	Below Sta	At/Near S	Below Standard		
No	9	White	Female	No	No	Yes	Demo dist	Demo Inst	Spring 202	Most Rece	04/06/202	04/06/202	826	10	Level 4	At/Near S	At/Near S	Above Standard		
No	9	White	Male	No	No	No	Demo dist	Demo Inst	Spring 202	Most Rece	04/04/202	04/04/202	821	8	Level 3	At/Near S	At/Near S	At/Near Standard		
No	9	White	Male	No	No	No	Demo dist	Demo Inst	Spring 202	Most Rece	04/04/202	04/06/202	803	8	Level 2	At/Near S	At/Near S	At/Near Standard		
No	9	White	Female	No	No	No	Demo dist	Demo Inst	Spring 202	Most Rece	04/05/202	04/07/202	812	9	Level 3	At/Near S	At/Near S	At/Near Standard		
No	9	White	Male	No	No	No	Demo dist	Demo Inst	Spring 202	Most Rece	05/09/202	05/09/202	801	9	Level 2	At/Near S	At/Near S	At/Near Standard		
No	9	White	Male	No	Yes	Yes	Demo dist	Demo Inst	Spring 202	Most Rece	04/05/202	04/05/202	800	9	Level 2	At/Near S	At/Near S	Below Standard		
No	9	White	Male	No	No	No	Demo dist	Demo Inst	Spring 202	Most Rece	04/04/202	04/04/202	847	9	Level 4	Above Sta	Above Sta	Above Standard		
No	9	White	Female	No	No	Yes	Demo dist	Demo Inst	Spring 202	Most Rece	04/05/202	04/07/202	834	9	Level 4	At/Near S	Above Sta	Above Standard		
No	9	White	Male	No	Yes	Yes	Demo dist	Demo Inst	Spring 202	Most Rece	04/05/202	04/05/202	787	9	Level 2	At/Near S	At/Near S	Below Standard		
No	9	White	Male	No	Yes	No	Demo dist	Demo Inst	Spring 202	Most Rece	04/04/202	04/04/202	808	8	Level 3	At/Near S	At/Near S	At/Near Standard		
No	9	White	Male	No	No	No	Demo dist	Demo Inst	Spring 202	Most Rece	05/15/202	05/22/202	804	9	Level 2	At/Near S	At/Near S	At/Near Standard		
No	9	White	Male	No	No	Yes	Demo dist	Demo Inst	Spring 202	Most Rece	04/04/202	04/06/202	829	9	Level 4	Above Sta	At/Near S	Above Standard		
No	9	White	Male	No	No	No	Demo dist	Demo Inst	Spring 202	Most Rece	04/05/202	04/05/202	802	9	Level 2	At/Near S	At/Near S	At/Near Standard		
No	9	Multi-Raci	Male	No	No	Yes	Demo dist	Demo Inst	Spring 202	Most Rece	04/05/202	04/05/202	788	9	Level 2	At/Near S	Below Sta	At/Near Standard		
No	9	White	Male	No	No	No	Demo dist	Demo Inst	Spring 202	Most Rece	04/04/202	04/04/202	802	8	Level 2	At/Near S	At/Near S	At/Near Standard		
Yes	9	White	Male	No	No	Yes	Demo dist	Demo Inst	Spring 202	Most Rece	04/05/202	04/05/202	806	9	Level 2	At/Near S	At/Near S	At/Near Standard		
No	9	White	Female	No	No	No	Demo dist	Demo Inst	Spring 202	Most Rece	04/05/202	04/07/202	816	8	Level 3	Above Sta	At/Near S	At/Near Standard		

1.5 Test Information Distribution Engine

Test Completion Rate Reports are available on the Test Information Distribution Engine (TIDE) website (https://mt.tide.cambiumast.com). These reports indicate the students who completed or need to complete computer-based testing and allow users to view participation summary statistics (counts and percentages) of students who have tested.

Once a user logs in, he or she is directed to the homepage, which allows users to access the Test Completion Rate Reports.

The Test Completion Rate Report allows teachers, principals, and district staff to see which students have not yet completed their tests. Users can select from a series of options to customize the group of students whose participation status is to be reviewed for a particular grade and subject, such as those who started but have not completed their test or those who have not yet begun their test. Users can export the list into a Microsoft Excel file and download the file.

1.6 Paper Individual Student Reports for Families

Individual Student Reports (ISRs) are printed by the district and distributed to parents. The primary purpose of the ISR is to provide a document that enables families to understand their child's performance in the subject in which he or she tested. The ISR also presents information that indicates how a student's performance compares to that of other students who took the same test. The report is organized as follows:

- **Top of Report.** The student's name, student ID, test grade, test date, school, and district are identified. Science reports include a frequently asked questions section here.
- **Montana Science Assessment Scores**. The student's scale score and corresponding performance level are displayed graphically and explained in accompanying text. A range of scores that is ±SEM is given with explanatory text.

- **Student Performance Compared.** Included with the Montana Science Assessment scores graphic, this section provides a comparison between the student's scale score and that of the student's school and district average scale scores.
- **Discipline Level Scores.** Discipline-level tables show how students performed on each discipline level. This section includes graphical displays of the Performance Category (*Below Mastery, At/Near Mastery, or Above Mastery*), or relative strength/weakness, for each of the discipline levels assessed. These results are explained in greater detail next to the graphics.

An example of the printed ISRs is displayed in Appendix 6-B, Sample Printed Individual Student Report.

2. Interpretation of Reported Scores

A student's performance on a test is reported as a scale score and a performance level for the overall test and as a performance level for each discipline level. Students' scores and performance levels are summarized at the aggregate levels. This section describes how to interpret these scores.

2.1 Scale Score

A scale score is used to describe how well a student performed on a test and can be interpreted as an estimate of a student's knowledge and skills as measured by their performance on the test. A scale score is the student's overall numeric score. These scores fall on a continuous scale. The Montana Science Assessment scale scores are not expressed on a vertical scale, which means that scores from different grades cannot be compared.

Scale scores can be used to illustrate a student's current levels of performance. Low scale scores indicate that a student does not possess sufficient knowledge and skills measured by the test. Conversely, high scale scores indicate that a student has proficient knowledge and skills measured by the test. When combined across a student population, scale scores can also describe school- and district-level changes in performance and reveal gaps in performance among different groups of students. In addition, scale scores can be averaged across groups of students, allowing educators to use group comparison. Interpretation of scale scores is more meaningful when the scale scores are used along with performance levels and Performance-Level Descriptors (PLDs). It should be noted that the utility of scale scores is limited when comparing smaller differences among scores (or averaged group scores), particularly when the difference among scores is within the standard error of measurement (SEM). Furthermore, the scale score of individual students should be cautiously interpreted when comparing two scale scores, because small differences in scores may not reflect real differences in performance.

2.2 Standard Error of Measurement

A student's score is best interpreted when recognizing that the student's knowledge and skills fall within a score range and are not just precise numbers. A scale score (the observed score on any test) is an estimate of the true score. If a student takes a similar test several times, the resulting scale scores will vary across administrations; sometimes the scores will be a little higher, a little lower, or the same. The SEM represents the precision of the scale score, or the range in which the student would likely score if a similar test were administered several times. The SEM can be interpreted as the degree of uncertainty of a student's score based on a statistical analysis of the student's answers on a test. When interpreting scale scores, it is recommended to always consider the range of scale scores incorporating the SEM of the scale score.

The "±" next to a student's scale score provides information about the certainty, or confidence, of the score's interpretation. The boundaries of the score band are one SEM above and below the student's observed scale score, representing a range of score values that is likely to contain the true score. For example, "680 ± 10" indicates that if a student were tested again, it is likely that he or she would receive a score between 670 and 690.

2.3 Performance Level

Performance levels are proficiency categories on a test, which students fall into based on their scale scores. For the Montana Science Assessment, scale scores are mapped into four performance levels (*Level 1*, *Level 2*, *Level 3*, *Level 4*) using performance standards (see Section 0, Student performance on each reporting discipline is reported in three achievement categories: *Below Standard, At/Near Standard, and Above Standard*. Unlike the achievement levels for the overall test, student performance on each of the discipline levels is evaluated with respect to the Level 3 achievement standard . Students performing at either *Below Standard* or *Above Standard* can be interpreted as having student achievement that is clearly below or above the Level 3 cut score for a specific discipline level. Discipline scores are scaled so that the same cut scores can be used for performance-level cuts. Students performing at *At/Near Standard* can be interpreted as having student achievement that does not provide enough information to tell whether students reached the Level 3 mark for the specific discipline level.

Cut Scores). PLDs are a description of content-area knowledge and skills that students at each performance level are expected to possess. Thus, performance levels can be interpreted based on PLDs. Students performing on the Montana Science Assessment at Level 3 and Level 4 are considered on track to demonstrate progress toward mastery of the knowledge and skills necessary for college and career readiness.

2.4 Performance Category for Discipline Levels

Student performance on each reporting discipline is reported in three achievement categories: *Below Standard, At/Near Standard, and Above Standard*. Unlike the achievement levels for the overall test, student performance on each of the discipline levels is evaluated with respect to the Level 3 achievement standard . Students performing at either *Below Standard* or *Above Standard* can be interpreted as having student achievement that is clearly below or above the Level 3 cut score for a specific discipline level. Discipline scores are scaled so that the same cut scores can be used for performance-level cuts. Students performing at *At/Near Standard* can be interpreted as having student achievement that does not provide enough information to tell whether students reached the Level 3 mark for the specific discipline level.

2.5 Cut Scores

For each performance level, there is a minimum and a maximum scale score that define the range of scale scores students in each performance level have achieved. Collectively, these minimum and maximum scale scores are defined as cut scores and are the cut-off points for each performance level. Table 4 presents the cut scores for science for all grades.

Table 4. Montana Science Assessment Science Performance-Level Cut Scores

Grade	Level 1	Level 2	Level 3	Level 4
5	400–476	477–505	506–530	531–600
8	700–780	781–807	808–825	826–900

2.6 Aggregated Scores

Students' scale scores are aggregated at the roster, teacher, school, and district levels to represent how a group of students performs on a test. When students' scale scores are aggregated, the aggregated scale scores can be interpreted as an estimate of knowledge and skills that a group of students possesses. This interpretation makes aggregated scores a powerful tool when comparing student performance across different groups of students, whether it be at a similar level of aggregation (e.g., school to school) or an analysis of a subgroup (e.g., comparing a teacher's roster to the overall school).

Given that student scale scores are estimates, the aggregated scale scores are also estimates and are subject to measures of uncertainty, as expressed using the calculated SEM for an aggregate average scale score. In addition to the aggregated scale scores, the percentage of students in each performance level is reported at the aggregate level to represent how well a group of students performs overall and by discipline level.

2.7 Relative Strengths and Weaknesses for Discipline Levels

For Disciplinary-Level performance, relative strengths and weaknesses at each standard are reported for aggregate levels only (e.g., classroom, school, district). Because an individual student responds to too few items within a standard to generate reliable data, the standard performance is produced by aggregating all items within a standard across students at an aggregate level.

The "Areas Where Performance Indicates Proficiency" section (the "Proficient?" column shown in Figure 4) for a standard shows how a group of students performed in each standard relative to the expected performance for proficiency. For summative tests, the expected performance for proficiency is the level of performance necessary to reach achievement Level 3. The reports in this section are standards-based reports in which the group performance in each standard is compared to the achievement standards for that standard. Similar to the achievement levels provided for the total test, the levels reported in this section indicate students' achievement with respect to the standards.

Since the “Areas Where Performance Indicates Proficiency” data for each standard is a comparison to the standards-based expectations, performance across groups can be compared.

For “Areas of Strongest and Weakest Performance” (the “Weak or Strong?” column shown in Figure 4) the expected performance is determined based on the students’ overall performance on the entire test. It shows how a group of students performed in each standard relative to their performance on the test overall. Rather than comparing across groups, “Areas of Strongest and Weakest Performance” provides more information regarding the relative strength and weakness on different standards on the test within a group.

2.8 Appropriate Uses for Scores and Reports

Assessment results can be used to provide information on individual student performance on the test. Overall, assessment results tell what a student knows and is able to do in certain subject areas and gives further information on whether a student is on track to demonstrate the knowledge and skills necessary for college and career readiness. Additionally, assessment results can be used to identify a student’s relative strengths and weaknesses in certain content areas. For example, performance categories for reporting disciplines can be used to identify an individual student’s relative strengths and weaknesses among reporting categories within a content area.

Assessment results on student performance on the test can be used to help teachers or schools make decisions on how to support students’ learning. Aggregate score reports at the teacher and school level provide information about the strengths and weaknesses of students and can be used to improve teaching and student learning. For example, a group of students may have performed very well overall, but possibly did not perform as well in several standards compared to their overall performance. In this case, teachers or schools can identify strengths and weaknesses of their students through the group performance by standards and promote instruction on specific areas where student performance is below their overall performance. Further, by narrowing down the student performance result by subgroup, teachers and schools can determine what strategies may be needed to improve teaching and student learning, particularly for students from disadvantaged subgroups. For example, teachers might see student assessment results by gender and observe that a particular group of students is struggling with Physical Sciences. Teachers can then provide additional instructions that focus on the Physical Sciences for these students.

In addition, assessment results can be used to compare student performance among different students and among different groups. Teachers can evaluate how their students perform compared with other students in schools and districts by overall scores and by discipline level. Although all students are administered different sets of items under the linear-on-the-fly test design, scale scores are comparable across students.

While assessment results provide valuable information to understand student performance, these scores and reports should be used with caution. It is important to note that scale scores are estimates of true scores and therefore do not represent the precise measure for student performance. A student’s scale score is associated with measurement error, and thus, users need to consider measurement error when using student scores to make decisions about student performance. Moreover, although student scores may be used to help make important decisions about student placement and retention or teachers’ instructional planning and implementation, the assessment results should not be used as the only source of information. Given that assessment results measured by a test provide limited information, other sources on student performance, such as classroom assessment and teacher evaluation, should be considered when making decisions on student learning. Finally, when student performance is compared across groups, users need to consider group size. The smaller the group, the larger the measurement error related to these aggregate data, thus requiring a more cautious interpretation.

3. SUMMARY

Montana Science Assessment results are reported online via the Montana Reporting System, as well as through printed Individual Student Reports (ISRs) sent to families. Results were released after the testing window closed and standard setting was completed.

The Reporting System is interactive. When educators or administrators log in, they see a summary of data about students for whom they are responsible (e.g., a principal will see the students in his or her school; a teacher will see students in his or her class). Users can then drill down through various levels of aggregation all the way to ISRs. The system allows them to tailor the content more precisely, moving from subject area through reporting categories, and even to standards-level reports for aggregates. Aggregate reports are available at every level, and authorized users can print these reports or download them (or the data on which they are based). ISRs can be produced individually or batched as PDF reports.

All authorized users can download files, including data about students for whom they are responsible, at any time. The various reports available may be used to inform stakeholders regarding student performance and instructional strategies.

Appendix 6-A
Reporting System User Guide

Montana Office of Public Instruction

Reporting System User Guide

For Summative and Interim Assessments

2022–2023

Published September 7, 2022

Prepared by Cambium Assessment, Inc.

Table of Contents

Introduction to the User Guide

This user guide gives instructions on using the Reporting System for the following:

- Accessing summative assessment data.
- Accessing interim assessment data.
- Scoring interim assessments.

It includes the following sections:

- How to Navigate Reports for Summatives and Interims
- How to Set Up Reports for Summatives and Interims to Suit Your Needs
- How to Export and Print Data on Summatives

The guide also includes additional information on accessing your interim data and scoring interims:

- How to Access Item-Level Data on Interims
- How to Score Items on Interims
- How to Set Up Interim Reports to Suit Your Needs
- How to Export and Print Data on Interims

How to Navigate Reports for Summatives and Interims

This section explains how to navigate your reports for both summative and interim assessments.

How to Understand Which Students Appear in Your Reports

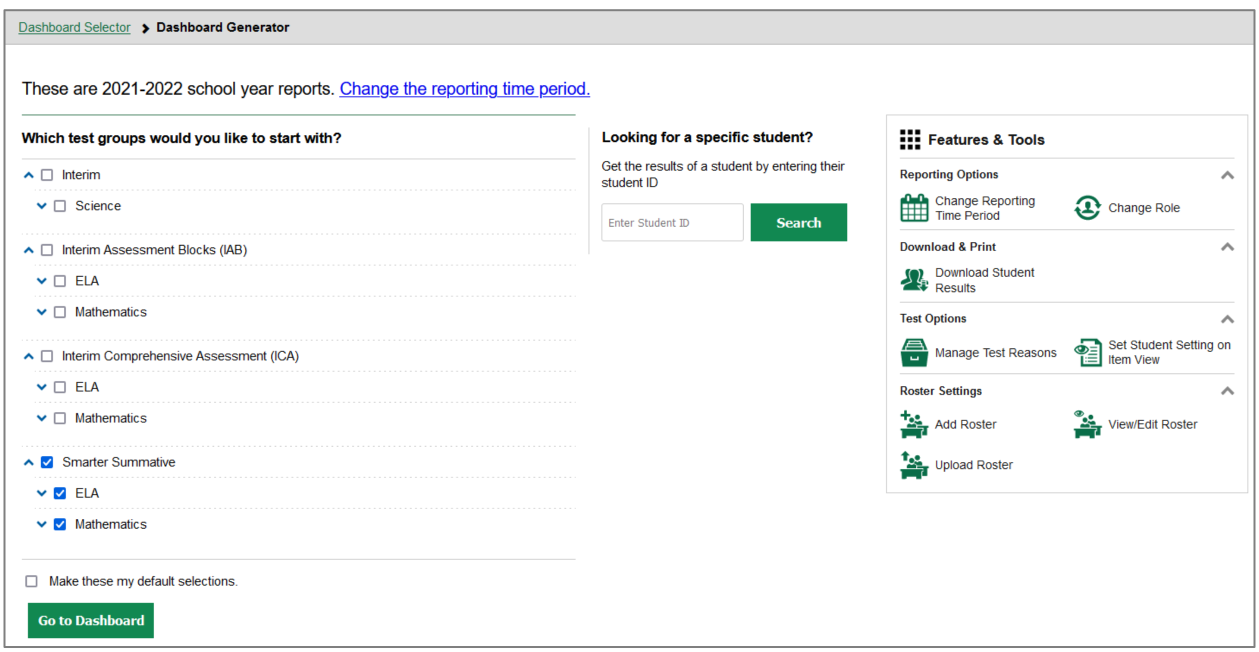
- Teachers can view data for all students in their classes (rosters) who have completed assessments. They can also view data for students to whom they have administered assessments in the current school year.
- School-level users can view data for all students in their schools who have completed assessments.
- District-level users can view data for all students in their districts who have completed assessments.

How to Start Viewing Reports

How to Use the Dashboard Generator Page

When you log in to the Reporting System, the Dashboard Generator page appears (Figure 1). The controls on the left let you select from the groups of tests that have been processed in your state for the school year. Typically, you may choose test types, subjects within the test types, and grades within the subjects.

Figure 1. Dashboard Generator Page



To the right of the dashboard generation controls are some other useful features. The *Looking for a specific student?* section allows you to use student ID to view all that student's test results. The **Features & Tools** menu ⠿ appears on the right side of this page and in reports and offers multiple features that are described elsewhere in this guide.

If the Dashboard Generator shows a message saying there are no data, that means no test opportunities are available for this school year.

1. *Optional*: To change the test groups listed in the Dashboard Generator, change the reporting time period.
2. Select the tests you want to view, expanding the test groups as needed.
3. *Optional*: To save your selections, mark the checkbox **Make these my default selections**. These selections will be set in the Dashboard Generator whenever you log in. You can change the defaults at any time.
4. Click **Go to Dashboard**. The dashboard appears, displaying any data available for your selections.

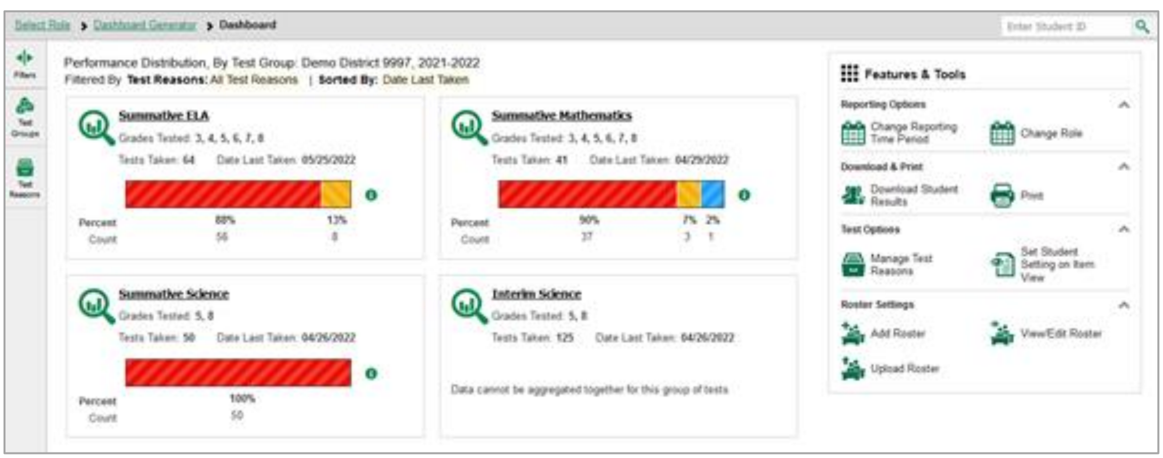
How to Use the Dashboard to View Aggregate Test Results

All users can view the standard dashboard (Figure 2). It displays aggregation cards representing test groups.

Each aggregation card displays the test group name, a list of grades, the number of students who took tests in the group, the date of the test last taken, and a performance distribution bar displaying both percentages and student counts below it. You may sometimes see the message “Data cannot be aggregated together for this group of tests” instead of the performance distribution bar for tests that do not report performance distribution, or that use different sets of performance levels. Test group cards are sorted by date last taken.

Clicking the button beside the performance distribution bar displays a legend with more information about performance levels.

Figure 2. Dashboard



If a message appears saying “There are no assessments to display,” you may not have any students who have taken tests in your selected test groups in the selected time period. You can change the reporting time period. If you are a teacher, you may also be able to view more students’ data by managing your classes (rosters).

To change the test groups and test reasons that appear, use the **Filters** panel on the left. For more information on filtering, see How to Set Up Reports for Summatives and Interims to Suit Your Needs and How to Set Up Interim Reports to Suit Your Needs. You can also backtrack to the Dashboard Generator using the link in the path at the upper-left corner and change your test group selections there.

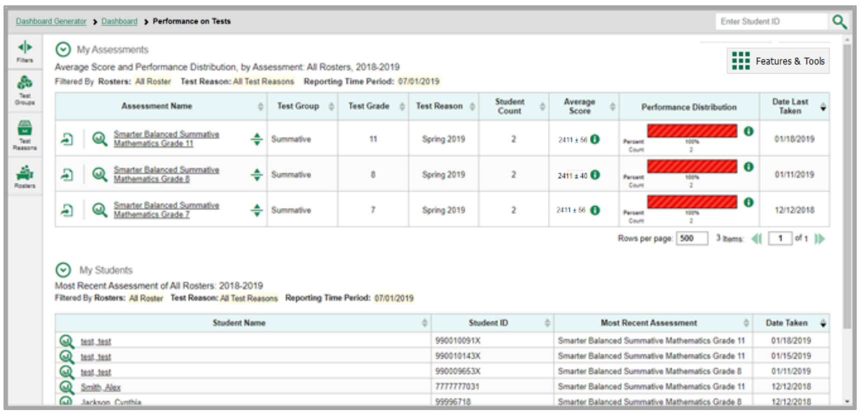
How to View More Detailed Data on a Particular Test Group

To view more detailed data for a particular test group, click the name of the group (or 🔍 beside it). The Performance on Tests report appears. It is filtered to display only the test group you selected.

In the Performance on Tests report, teachers see two tables, as in Figure 3:

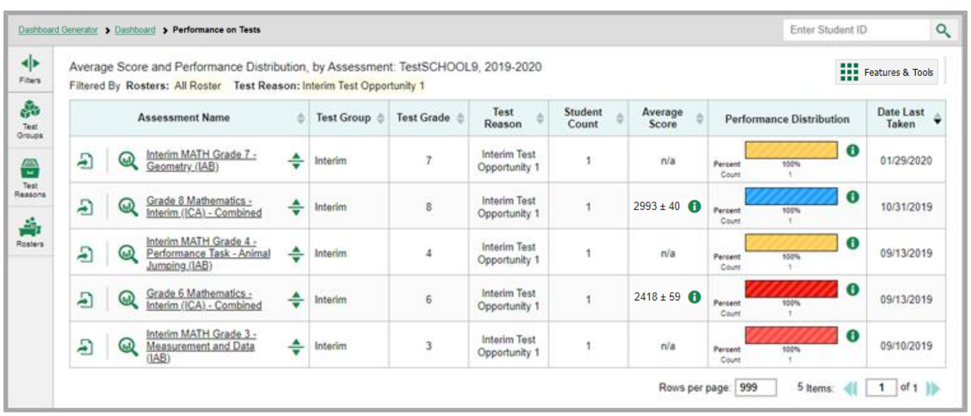
- The My Assessments table, listing all the assessments in the filtered test group or groups.
- The My Students table, listing all your students who took the assessments.

Figure 3. Teacher View: Performance on Tests Report



District- and school-level users see just one table, as in Figure 4. Like the first table on the teacher Performance on Tests report, this table lists all the assessments in the test group.

Figure 4. School-Level User View: Performance on Tests Report



For each test, the assessments table shows the test group, grade, test reason (the name of the test window of a summative assessment, or a category assigned to an interim assessment), number of students who took the test, average score, performance distribution, and date the test was last taken.

You can use the filters to view a different set of assessments. For more information on filtering, see How to Set Up Reports for Summatives and Interims to Suit Your Needs and How to Set Up Interim Reports to Suit Your Needs.

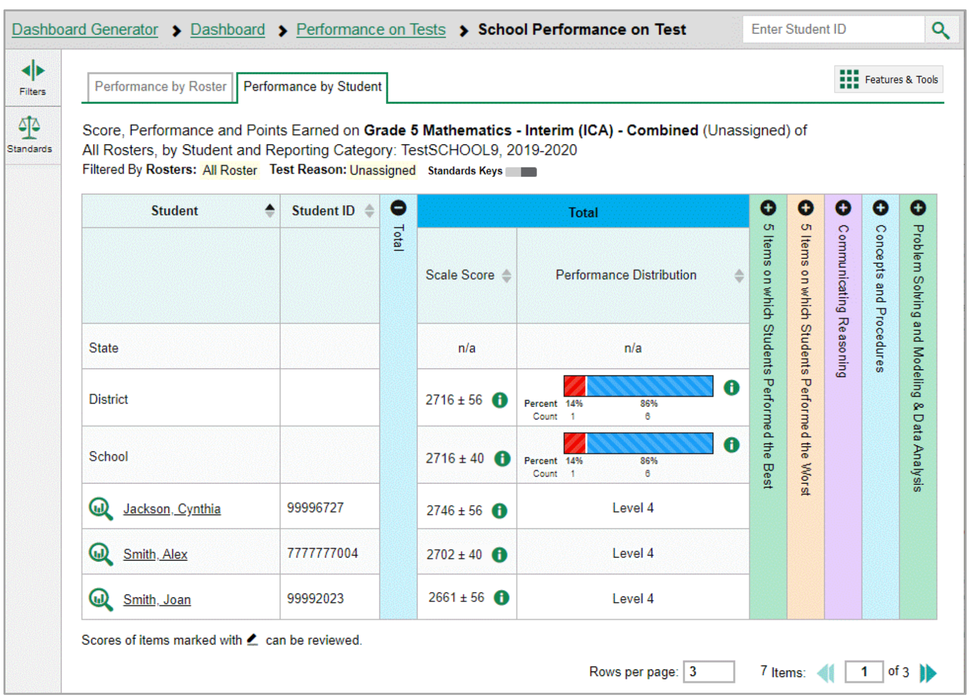
If a message appears saying “There are no assessments to display” or “There are no students to display”, you may have filtered out all data.

For Teachers and School-Level Users: How to View Student-Level Data for All Your Students

The **Performance by Student** tab displays test results for each of your students across classes (rosters). In order to see the results for all your students, follow the instructions below.

1. Starting from the dashboard that appears when you log in, click a test group name (or beside it).
2. Click a test name (or beside it) in the assessments table at the top of the page.
3. In the report that appears, select the **Performance by Student** tab, as in Figure 5. You will see results listing all your students. The first few rows also show aggregate performance data for your state, district, school, and/or total students.

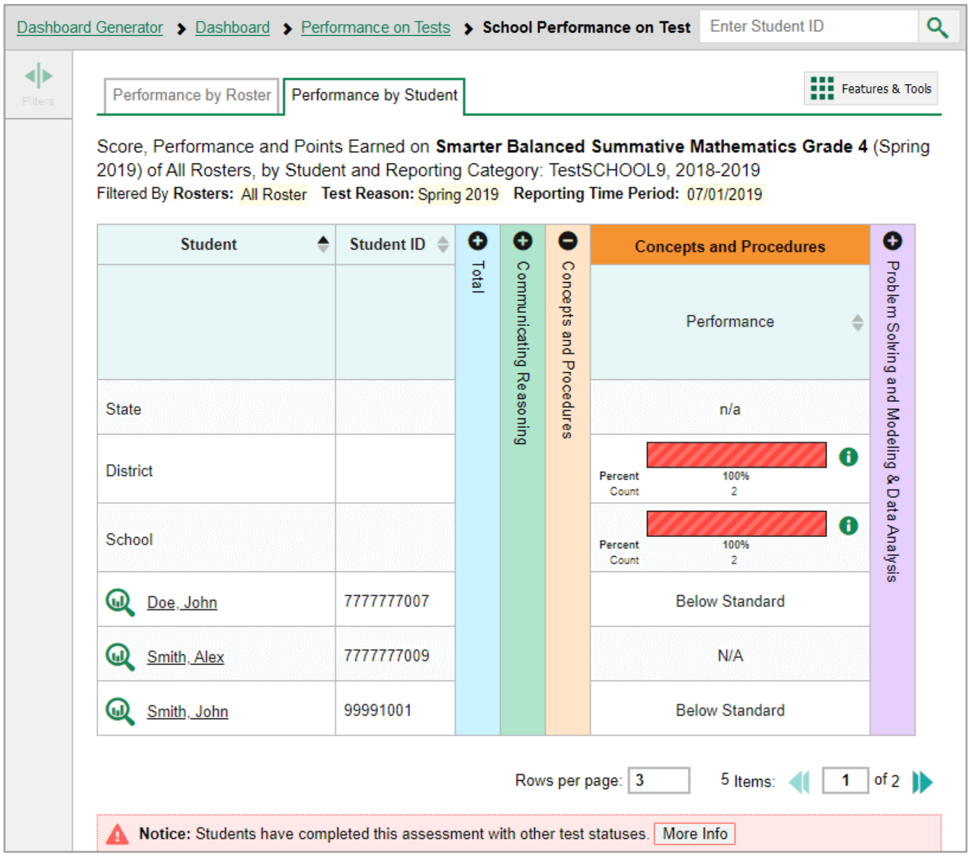
Figure 5. School Performance on Test Report: Performance by Student Tab



To see which students performed best, click the score or Performance columns to sort them.

You can view your students' performance in each area of the test by clicking the reporting category section bars to expand them, as shown in Figure 6.

Figure 6. My Students' Performance on Test Report: Performance by Student Tab with Expanded Reporting Category Section



How to View Test Results for Classes (Rosters) on a Particular Test

You can view a list of classes (rosters) that took a particular test, and you can also view the test results for a particular class.

How to Access Test Results for All Your Classes (Rosters)

The **Performance by Roster** tab (Figure 7) displays test results for each class (roster). To view this tab, follow the instructions for your user role below.

Teachers and school-level users:

1. Generate a dashboard and click a test group name (or beside it).

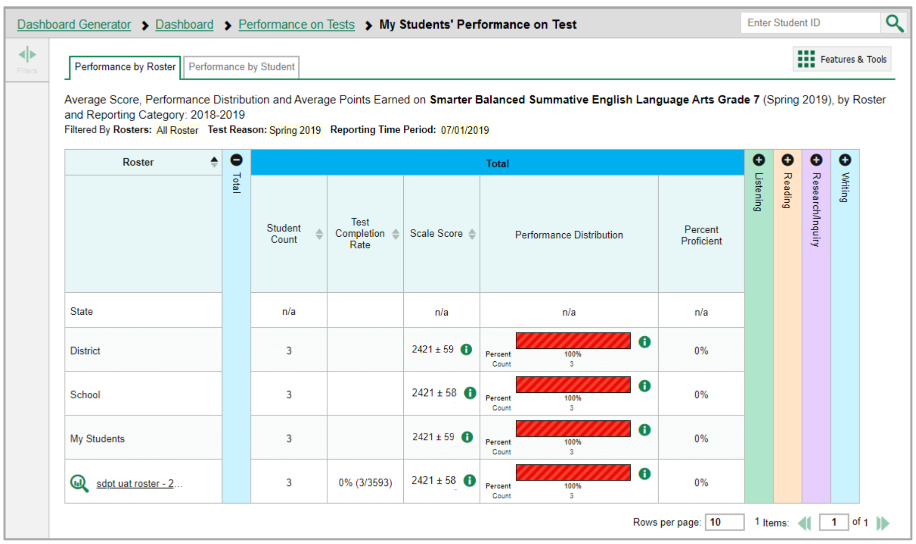
2. Click a test name (or beside it) in the table of assessments. Either the My Students' Performance on Test or the School Performance on Test report appears, depending on your role. It is open to the **Performance by Roster** tab.

District-level users can view all classes (rosters) in a school. To do so, follow these instructions:

1. Generate a dashboard and click a test group name (or beside it).
2. Click a test name (or beside it) in the table of assessments. A page of district test results appears, listing schools within the district.
3. Click a school name (or beside it). The School Performance on Test report appears, open to the **Performance by Roster** tab.

The report shown in Figure 7 displays a list of your classes (rosters) and each class's performance. The first few rows also show aggregate performance data for your state, district, school, and total students.

Figure 7. My Students' Performance on Test Report: Performance by Roster Tab



How to See Which Classes (Rosters) Performed Well on This Assessment

To see which classes performed best on the test, do either of these things:

- Click the score column header to sort by score and look for rosters with high average scores.
- Look at the bars in the Performance Distribution column to see where the percentage of students at or above proficient is high.

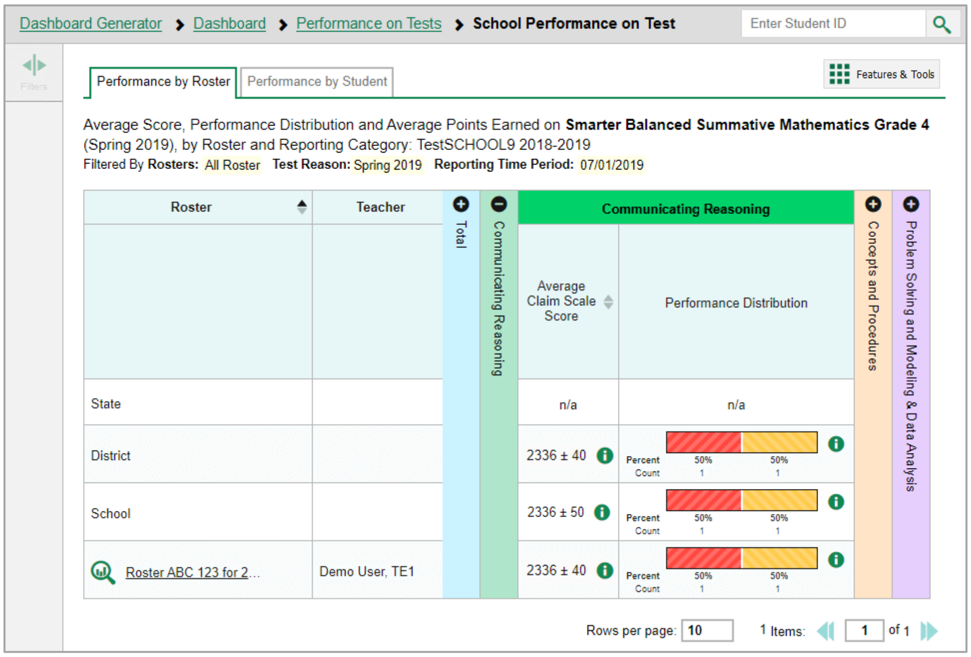
How to See Which Classes (Rosters) Had the Highest Test Completion Rates

To see which classes had the highest test completion rate, click the Test Completion Rate column header to sort the column.

How to See How Well Classes (Rosters) Performed in Each Area on the Test

For tests with reporting category sections, you can compare the performance of your students in each area of the test. Click each vertical section bar to expand or collapse it. In this example (Figure 8), you can view average score and a performance distribution bar for each class (roster) under the reporting category Communicating Reasoning.

Figure 8. My Students' Performance on Test Report: Performance by Roster Tab with Expanded Reporting Category Section



Summative and interim tests cover multiple reporting categories, while a benchmark test covers only one.

How to View and Interpret Standard Measures Within Reporting Category Sections

Aggregate test results for Smarter Balanced and Montana Science Assessment (MSA) tests may include detailed performance measures within standards within reporting categories.

An educational standard, sometimes called an assessment target, describes the skill the item measures. Standards are nested within clusters (groups of standards).

Note: The Alternate Montana Science Assessment (AMSA) tests are reported at the overall scale score only. No sub-scores or reporting categories are included.

The Standards sub-section (shown in Figure 9) contains the following:

- **Clusters** within the reporting category.
  - **Standards** within each cluster.

- – **Measures** within each standard.

To learn more about each standard, click the more information button to the right of the standard name.

Figure 9. School Performance on Test Report: Performance by Roster Tab with Expanded Reporting Category Section

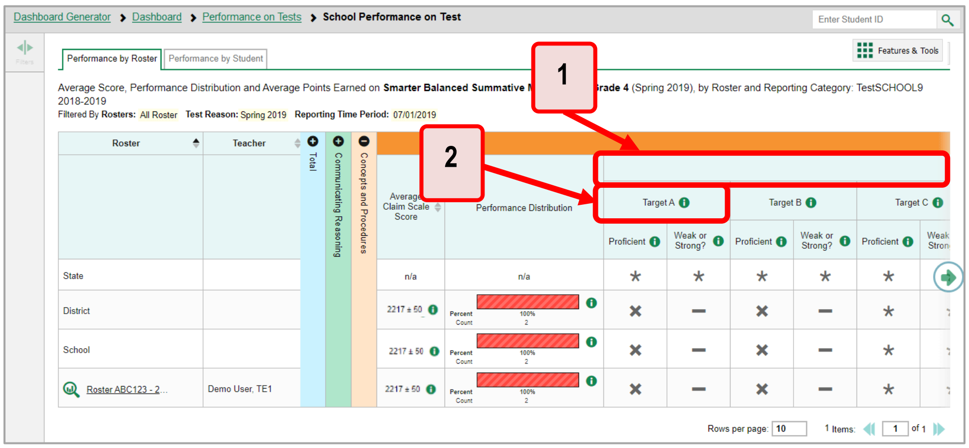


Table 1. School Performance on Test Report: Performance by Roster Tab: Reporting Category Section Elements

#	SBAC Elements	MSA Elements
1	Claim	Discipline
2	Targets	Disciplinary Core Idea (DCI)

Each standard column consists of sub-columns displaying performance measures:

- **Weak or Strong?:** You may want to identify strengths and weaknesses in different standards within the reporting category so you can focus instruction on very specific areas where improvement is needed. This column displays icons indicating how students' average performance on the standard compares with their average performance on the overall test. Note that these icons indicate only relative performance and not proficiency.
- **Proficient:** This column displays icons indicating whether students have on average attained proficiency in the standard.

To learn more about these measures and the symbols they use, click the more information button to the right of each measure.

How to View and Interpret Writing Dimension Measures (ELA ICA Only)

Aggregate test results for some assessments may also include a **Writing Dimensions** section to the right of the expandable sections in the report table. You can expand it by clicking the vertical bar, just as with the reporting category sections. This section helps you understand how students performed on different aspects of writing.

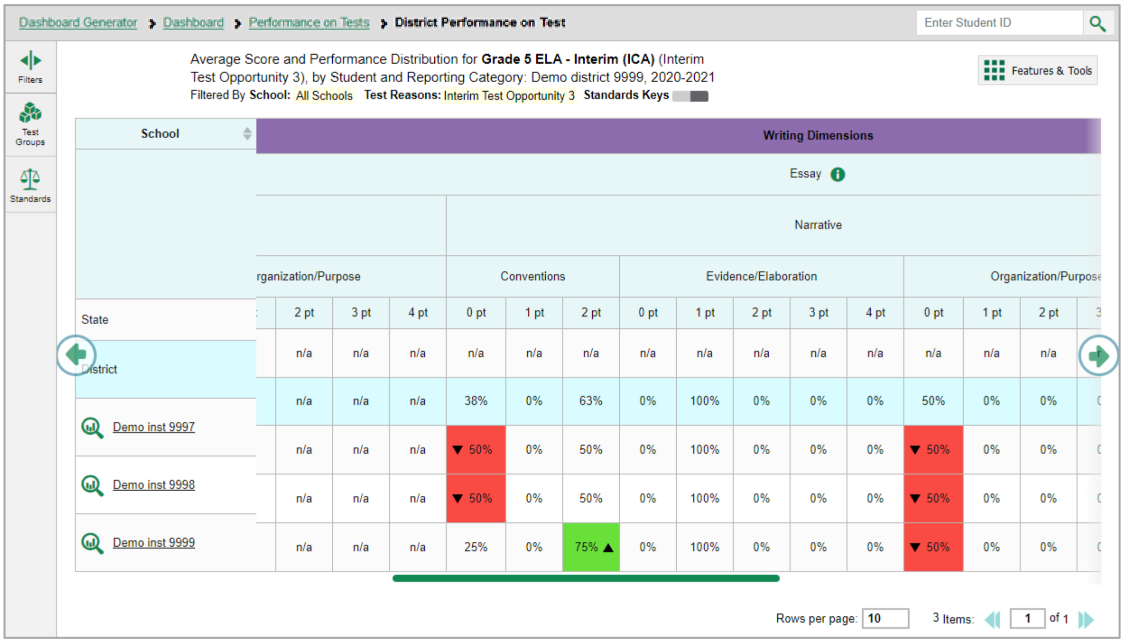
The **Writing Dimensions** section (see Figure 10) contains the following under the Essay header:

- **Essay type**. For example, Informative/Explanatory, Opinion, and Narrative.
  - **Writing dimension** within the essay type, as listed in item rubrics. For example, Purpose, Focus, and Organization; Evidence and Elaboration; and Conventions of Standard English.
    - **Point value**. A sub-column for each possible item point value for the writing dimension. Each point value sub-column displays the percentage of students who earned that number of points.

    For each dimension, the lowest ▼ and highest ▲ point values are sometimes highlighted and marked with arrow icons in the rows with the highest percentages. This allows you to quickly identify groups of students who are performing well and those who may need additional support.

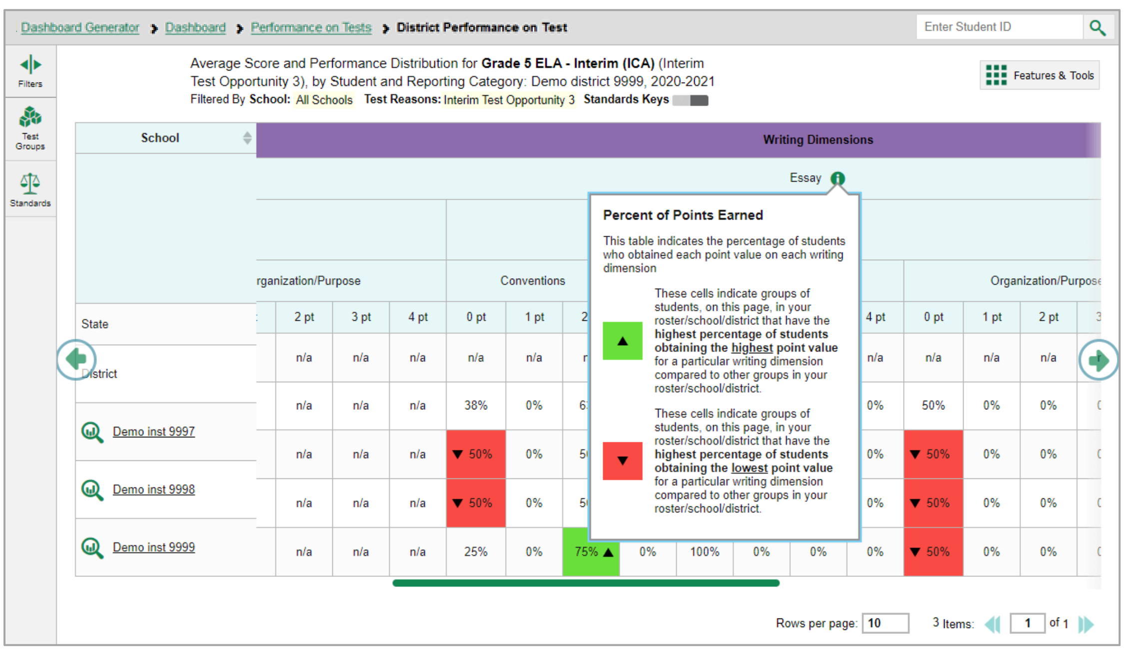
    Note that highlighting and arrow icons do not appear where they are not useful. A report containing only one row of data does not have them, and neither does a row in which the percentages are "n/a" or are all the same.

Figure 10. District Performance on Test Report with Expanded Writing Dimensions Section



To learn about the highlighting and arrow icons, click the more information button in the Essay header. A legend expands, as in [Figure 11](#).

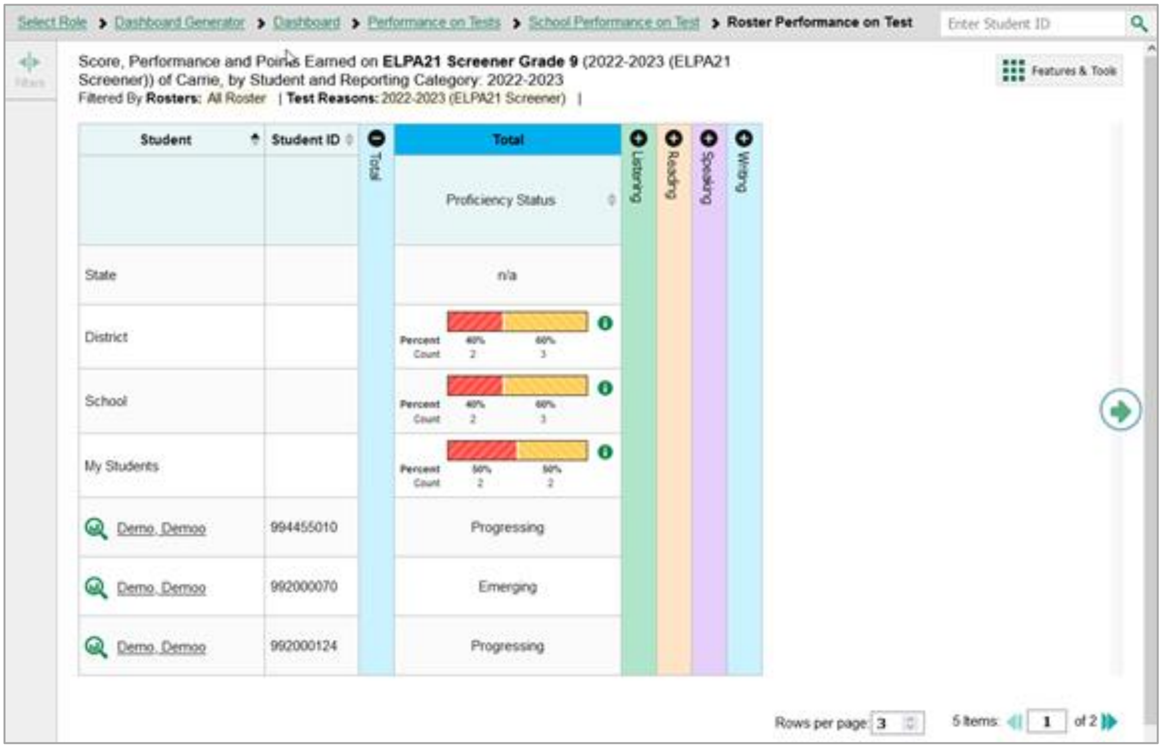
Figure 11. District Performance on Test Report with Expanded Writing Dimensions Section and Expanded Legend



How to Access Test Results for an Individual Class (Roster)

Prior sections explained how to access test results for all your classes (rosters). To view results for one specific class, click the name of a class that appears in the first column of the report (or beside it). The class results listed by student appear (see Figure 12).

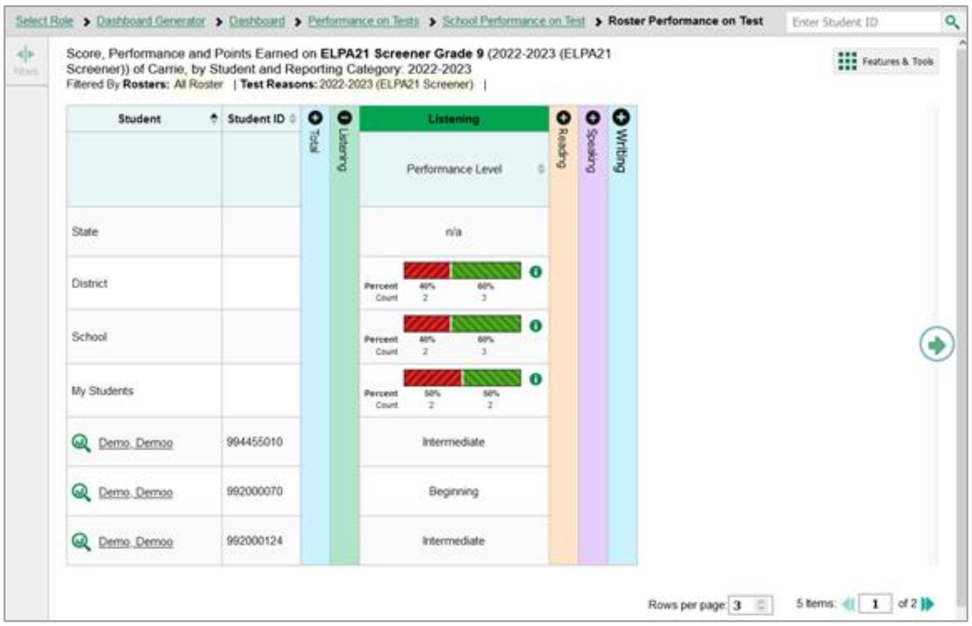
Figure 12. School-Level User View: Roster Performance on Test Report



How to See How Well Students in Your Class (Roster) Performed in Each Area of the Test

You can compare the performance of your students in each area of the test using the reporting category sections, as shown in Figure 13. Click the vertical section bar to expand each section.

Figure 13. School-Level User View: Roster Performance on Test Report with Expanded Reporting Category Section



For School- and District-Level Users: How to View Test Results for a School on a Particular Test

You can view test results for all the students in a school on a particular test. This gives you a high-level look at how the school is performing.

How to Access Test Results for a School

School-level users:

1. Generate a dashboard and click a test group name (or beside it).
2. Click a test name (or beside it) in the table of assessments. The School Performance on Test report appears.

District-level users:

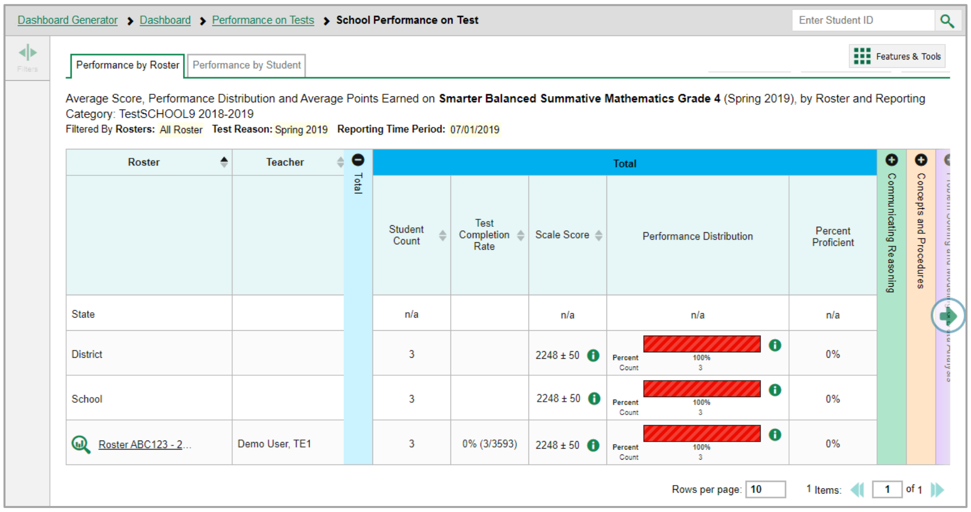
1. Generate a dashboard and click a test group name (or beside it).
2. Click a test name (or beside it) in the table of assessments. A table listing test results by school appears.

3. Click the name of the school (or beside it) for which you would like to see results. The test results for the school appear.

   Note that district-level users can also access the test results from a school directly from the Performance on Tests report, by first filtering by school.

The **Performance by Roster** tab is open by default, as in Figure 14.

Figure 14. School Performance on Test Report: Performance by Roster Tab



How to See Which Classes (Rosters) Performed Well on This Assessment

In the **Performance by Roster** tab, look at the Performance Distribution column and click the header of the score column to sort by score. Rosters with a high average scale score, and with a high percentage of students performing at or above proficient in the performance distribution bar, performed well on the assessment. If certain classes (rosters) performed consistently well, you could use them as a model for the classes with lower performance.

For District-Level Users: How to View Test Results for a District on a Particular Test

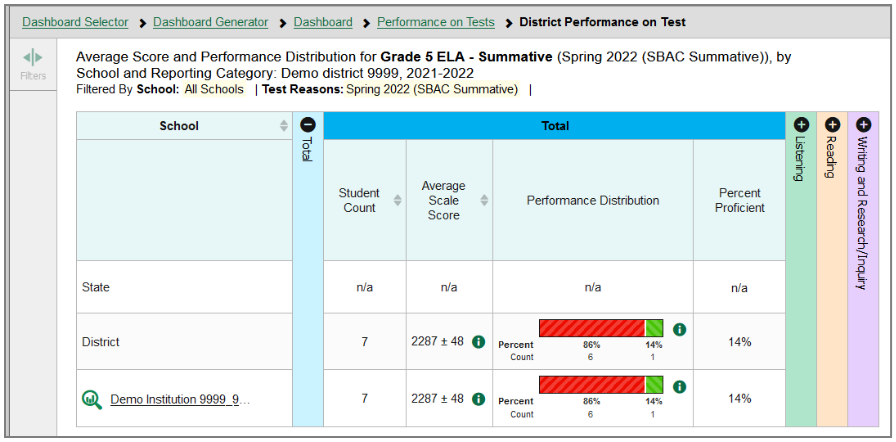
You can view test results for a district on a particular test. This gives you a high-level view of how the district is performing.

How to Access Test Results for a District

1. Generate a dashboard and click a test group name (or beside it).

2. Click a test name (or beside it) in the table of assessments. The District Performance on Test report appears, listing schools in the district (see Figure 15).

Figure 15. District Performance on Test Report



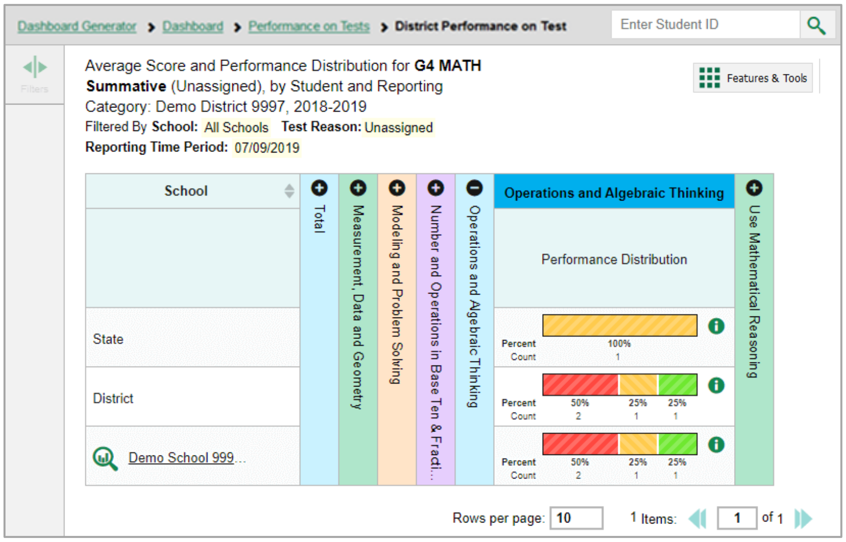
How to See Which Schools in the District Performed Well on This Assessment

Look at the score column and/or Performance Distribution column, and click the score column to sort by it.

How to See How Well Schools in the District Performed in Each Area of the Test

Click the vertical section bars to expand the reporting category sections (as in Figure 16).

Figure 16. District Performance on Test Report with Expanded Reporting Category Section



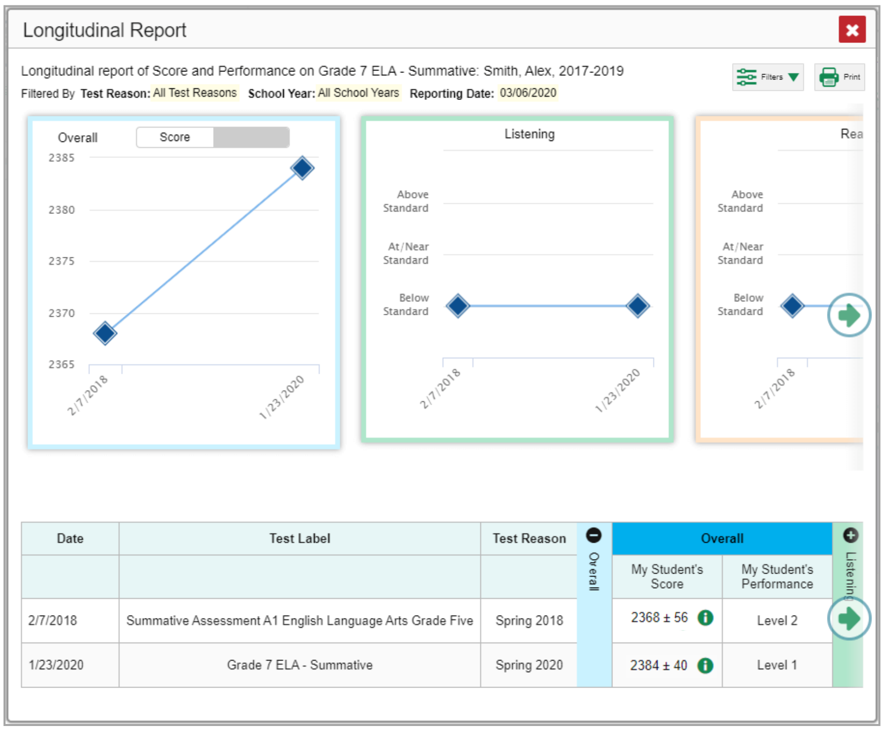
How to Track Student Performance Over Time

You can view your students' performance over time across multiple related assessments or across multiple test opportunities of a single assessment. This lets you see how students' performance has improved or declined.

Each Longitudinal Report displays performance data for one of the following:

- A group of students who have completed every assessment available in the report. If you are a district- or school-level user, note that a certain percentage of students must have taken *all* the related assessments in order for you to generate a Longitudinal Report. Teachers have the option of adjusting the pool of students, tests, and test reasons.
- An individual student (see Figure 17).

Figure 17. Longitudinal Report Window: Summative Report for a Single Student with Multiple Reporting Categories



How to Access a Longitudinal Report Comparing Related Assessments

If the student(s) in your test results have completed multiple related assessments, the **Build Longitudinal Report** button allows you to access a Longitudinal Report in the reports for any of those assessments. If they haven't done so, then no Longitudinal Report is available.

Click the **Build Longitudinal Report** button in the **Features & Tools** menu.

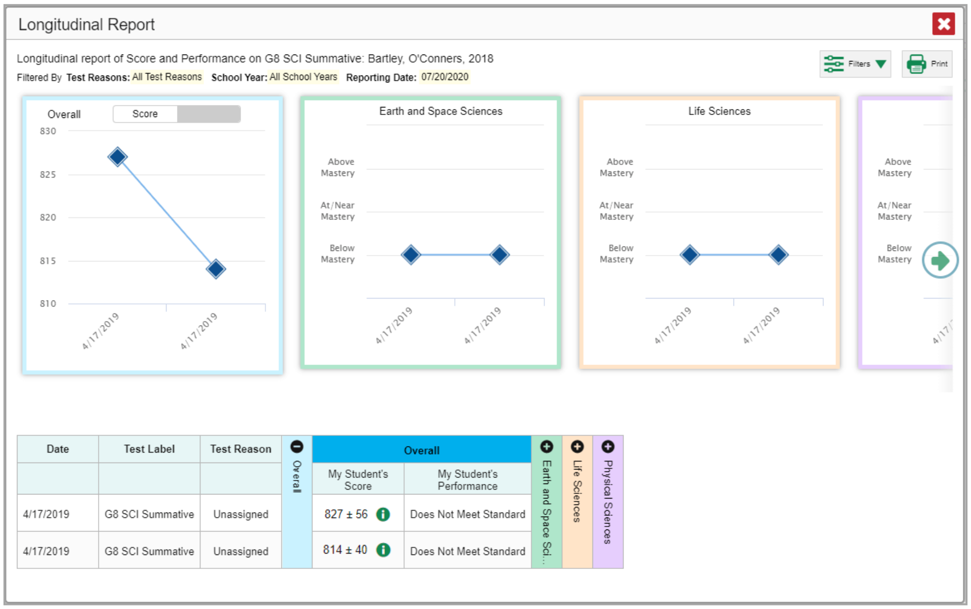
The ***Longitudinal Report*** window appears. Depending on your role, the test types, and the number of students in the report, it may display a report options page rather than the Longitudinal Report itself. The contents of this page depend on your role and the number of students.

How to View Students' Overall Performance on These Assessments Over Time

Look at the graph in the upper-left corner of the Longitudinal Report (see Figure 18). It shows the scores or performance levels of the student(s) each time they took the test(s).

Score data are plotted along a line.

Figure 18. Longitudinal Report Window: Summative Report for a Single Student with Multiple Reporting Categories



Performance level data are shown either the same way or, for multiple students, in performance distribution bars.

Mouse over the data points in a line graph or the sections in a bar to get more information.

Alternatively, in the table at the bottom of the report, look at the **Overall** section.

How to Switch Between Score Data and Performance Level Data

When a graph offers both score and performance level data, a toggle bar appears at the top of it. Click the toggle to switch. You may want to do this if you find performance level data easier to read, or if you prefer the precision of score data. Sometimes a test includes only one type of data.

How to See Students' Performance in Different Areas Over Time

Look at the reporting category graphs to the right of the overall performance graph, or look at the expandable reporting category sections in the table at the bottom. Here, you can see at a glance how students are improving or declining in each area, and you can compare their trajectories in different areas.

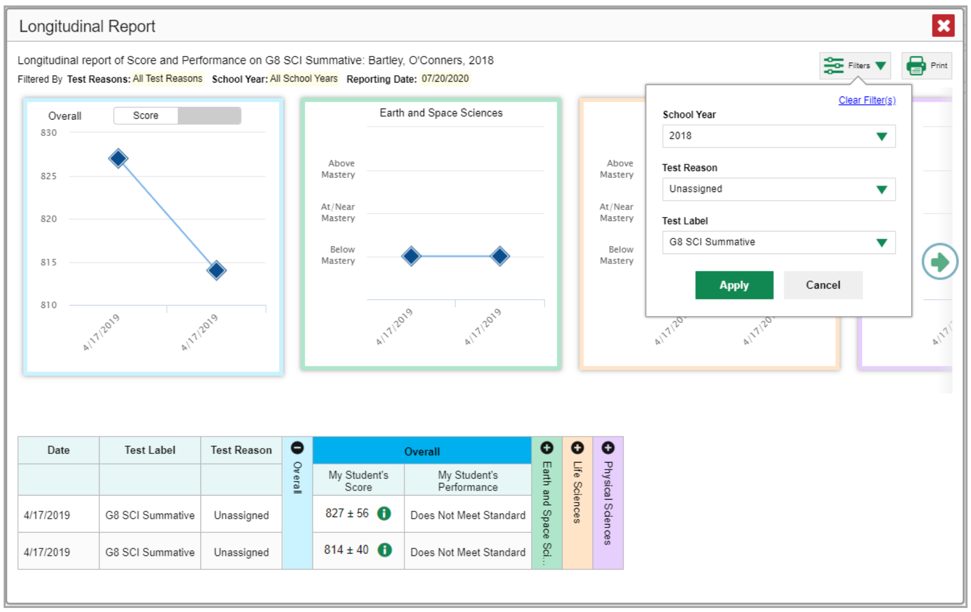
How to Filter Test Opportunities to Show Some and Not Others

You may want to filter a Longitudinal Report in order to focus on some test opportunities and not others.

Note that filtering tests may affect the set of students whose data are included in the report.

1. Open the **Filters** menu at the upper-right corner and select the filter options you prefer from the drop-down lists (see Figure 19).

Figure 19. Longitudinal Report Window: Summative Report for a Single Student with Multiple Reporting Categories and with Expanded Filter Menu



- You may want to filter by a particular school year or years. Note that years are not calendar years. "2021" refers to the 2021–2022 school year. By default, Longitudinal Reports show data for all years.

  Longitudinal Reports can show student performance from a time when the students were not yet associated with you. For example, if you are a seventh-grade teacher, you can use these reports to view your current students' performance on last year's sixth-grade tests.

- If the report includes interim assessments, you may wish to filter by a test reason (a category of test), which means excluding all other test reasons from the data. For example, you may want to narrow the report down to show only tests taken in the spring. For summative assessments, test reasons are the same as test windows and are not useful.

- Finally, you may find that certain individual tests are less relevant than others. In that case, you can use the **Test Label** options to deselect the names of the tests you don't want to see.

2. Click **Apply**.

3. *Optional*: To revert all filters to their defaults, open the **Filters** menu again and click **Clear Filters**. Click **Apply**.

A row of filter details appears below the report header, showing the test reasons and school years included in the report.

How to View Test Results Broken Down by Demographic Sub-Groups

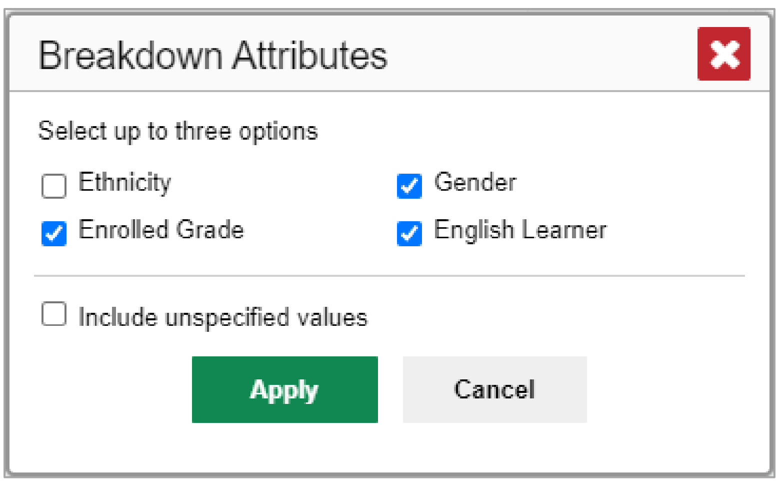
You can use the **Breakdown By** button in the **Features & Tools** menu to compare performance between different demographic sub-groups. This button is available for most aggregate test results.

How to View Test Results Broken Down by Demographic Sub-Groups

To view test results broken down by demographic sub-groups, do the following:

1. Click **Breakdown By** in the **Features & Tools** menu. The ***Breakdown Attributes*** window opens (see Figure 20).

Figure 20. My Students' Performance on Test Report: Performance by Student Tab: Breakdown Attributes Window



2. Select up to three student demographic categories.

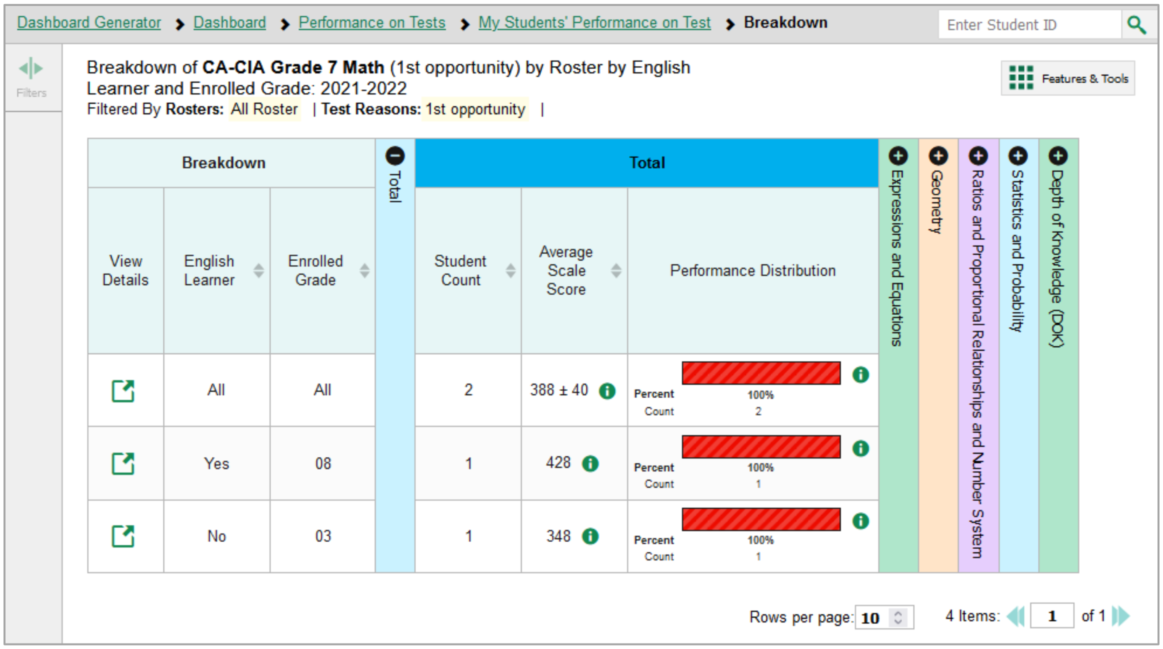
   There is also an **Include unspecified values** checkbox, explained below:

   - Some students who complete tests do not have specific demographic information in the Test Information Distribution Engine (TIDE). These students are considered to have unspecified values.
   - To include data for these students, mark the checkbox.

3. Click **Apply**.

Data for each sub-group selected are displayed in the report (see Figure 21).

Figure 21. Demographic Breakdown of a My Students' Performance on Test Report

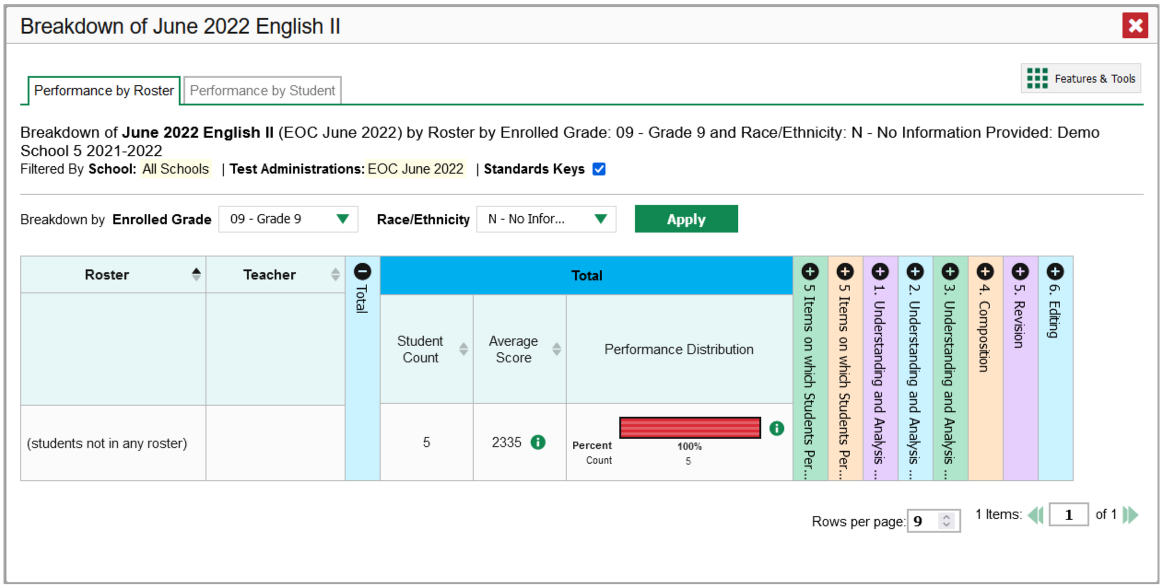


How to View Test Results for a Particular Demographic Sub-Group or Combination

When viewing test results broken down by sub-groups, go to the row for the demographic combination you want to view and click the view button in the View Details column on the left (see Figure 21).

A window opens, displaying detailed results for that combination. The report table is now laid out the same way as the original report, before you viewed it broken down by sub-groups (see Figure 22).

Figure 22. Demographic Combination Breakdown Window (from School Performance on Test Report)



At the top of the report table are filter menus for each demographic category you chose. To change the demographic combination displayed, use the filters to select the demographic sub-groups you want to see and click **Apply**. The new combination is displayed.

You can use this window to get an in-depth look at specific groups of students. For example, you may want to determine which classes (rosters) have the highest-performing girls in the first grade.

How to View Test Results for Individual Students

You can find out how well an individual student understands the material covered on a specific completed assessment. You can also view a report for all the assessments a student has taken.

How to Access Test Results for an Individual Student on a Particular Test

Teachers and school-level users:

1. Generate a dashboard and click a test group name (or beside it).
2. Click a test name (or beside it) in the table of assessments. A page of test results appears.
3. Select the **Performance by Student** tab.
4. Click the name of an individual student (or beside it) in the report. The Student Performance on Test report appears (see Figure 23).

District-level users:

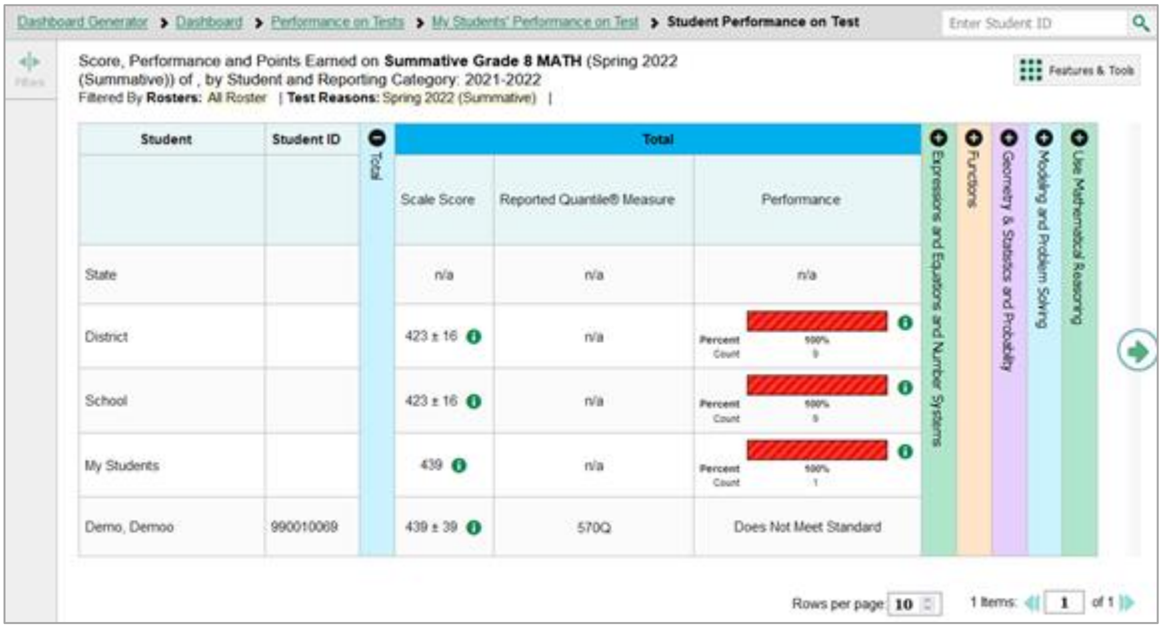
1. Generate a dashboard and click a test group name (or beside it).

2. Click a test name (or beside it) in the table of assessments. A page of test results by school appears.

3. Click a school name (or beside it). The School Performance on Test report appears.

4. Perform the same steps as teachers and school-level users, starting at step 3.

Figure 23. Teacher View: Student Performance on Test Report



You can view the student's performance in each area of the test using the reporting category sections, which you can click to expand.

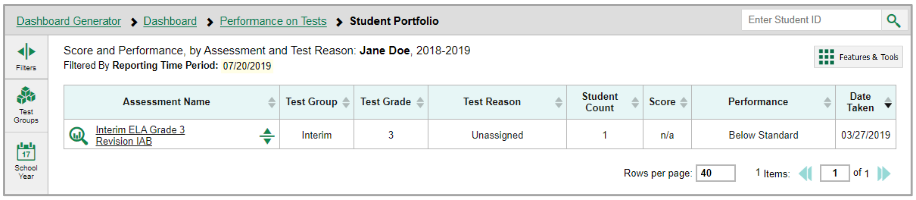
How to View a Report for All the Assessments a Student Has Completed Over Time

The Student Portfolio Report allows you to view all the assessments an individual student has completed over time. This is useful for viewing performance on tests that have multiple opportunities, and for interim tests that were administered multiple times throughout the year.

To access this report, enter the student's SSID in the search field in the upper-right corner and click . (When you are viewing the Dashboard Generator and data are available, the same field appears to the right of the dashboard generation controls.) The Student Portfolio Report appears (see Figure 24).

Teachers can also access this report from the Performance on Tests report by going to the My Students table below the main assessments table and clicking a student's name (or beside it).

Figure 24. Student Portfolio Report



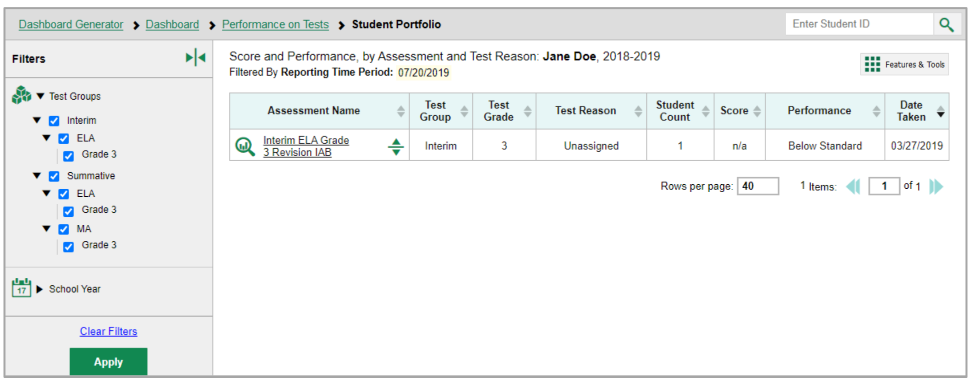
To view individual test results for this student, click a test name (or beside it).

How to Use the Student Portfolio Report to View Only the Tests You're Interested In

You can temporarily filter which tests you want to see in the Student Portfolio Report. You may want to do this, for example, if you are an ELA teacher and you don't want to see a student's math scores. By default, the data for those math assessments appear in the report, but you can exclude them.

1. In the **Filters** panel on the left side of the Student Portfolio Report, click either the expand button or the **Test Group** button . The **Filters** panel expands (see Figure 25).

Figure 25. Student Portfolio Report with Expanded Filters Panel



2. Mark as many selections as you like in the **Test Group** section of the **Filters** panel. Tests are organized by test type, subject, and grade.

3. Click **Apply**. The Student Portfolio Report updates to show only data for those tests.

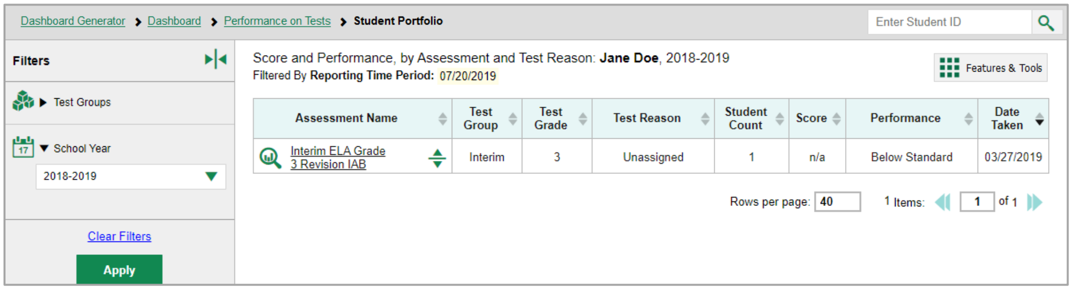
4. *Optional*: To revert all filters to their defaults, open the filters panel again and click **Clear Filters**. Click **Apply**. Filters also revert when you log out, switch user roles, or switch systems.

How to View a Student's Performance on Tests Taken in a Previous School Year or Years

If there are multiple years of data for a student, the Student Portfolio Report can look back at previous years. This gives you a high-level look at the student's progress. Student Portfolio Reports can show student performance from a time when the students were not yet associated with you. For example, if you are a seventh-grade teacher, you can use these reports to view a current student's performance on last year's sixth-grade tests.

1. On the left side of the page, click either the **Filters** panel expand button or the school year button. The **Filters** panel expands.
2. Under **School Year**, select a year or years (see Figure 26).

Figure 26. Student Portfolio Report with Expanded Filters Panel



3. Click **Apply**.

To switch back to the current year:

1. Open the **Filters** panel again.
2. Click **Clear Filters**.
3. Click **Apply**.

How to Use Individual Student Reports (ISR)

This section discusses Individual Student Reports (ISR), each of which provides easy-to-read performance data on a student's test. If a student took tests multiple times, an ISR will be available for each test opportunity (an instance of a test the student took). ISRs are useful for sharing performance information with students and their parents and guardians, and may be generated in the language of your choice.

What an Individual Student Report (ISR) Looks Like and How to Read It

An ISR is a PDF that shows results for a test opportunity. It may consist of a single page or multiple pages. ISR layouts vary according to the type of test. Details of sample ISRs are shown below in Figure 27, Figure 28, Figure 29, Figure 30, and Figure 31.

- At the top of each ISR are the student name and SSID, the name of the test, district, and school, and any other relevant information.
- Below that is a summary of the student's performance. An ISR for a scale-scored test displays the student's performance on a vertical scale that includes all the cut scores and performance levels.
- Each ISR includes a comparison table showing the average performance of the state, district, and/or school.
- Many ISRs include a table detailing the student's performance in each reporting category (as in Figure 28).
- Some ISRs include item-level data (as in Figure 29), scoring assertions, and/or scoring assertion outcomes (as in Figure 30).
- Some ISRs include longitudinal graphs (as in Figure 31).

Figure 27. Detail of Individual Student Report (ISR): Math Summative

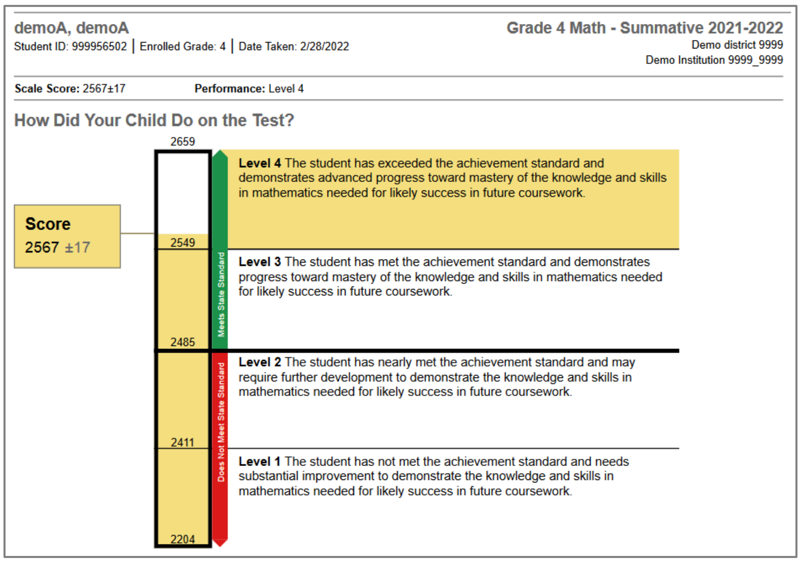


Figure 28. Detail of Individual Student Report (ISR): Math Interim (ICA) with Reporting Categories

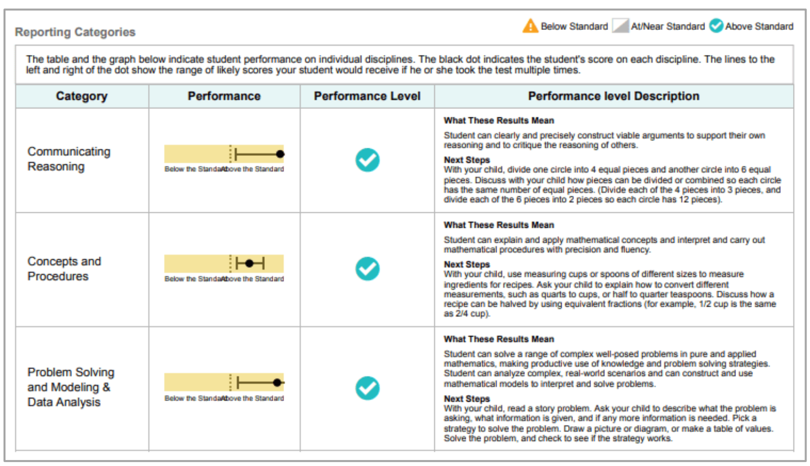


Figure 29. Detail of Individual Student Report (ISR): Math Interim (ICA) with Item- and Standard-Level Data

Communicating Reasoning		
Question #	**Standard**	**Points Earned/Points Possible**
3	Base arguments on concrete referents such as objects, drawings, diagrams, and actions.	0/2
5	Distinguish correct logic or reasoning from that which is flawed, and—if there is a flaw in the argument—explain what it is.	1/1
12	Test propositions or conjectures with specific examples.	0/1
17	State logical assumptions being used.	1/1
18	Use the technique of breaking an argument into cases.	1/1
28	Distinguish correct logic or reasoning from that which is flawed, and—if there is a flaw in the argument—explain what it is.	0/1
36	Base arguments on concrete referents such as objects, drawings, diagrams, and actions.	1/1
37	Base arguments on concrete referents such as objects, drawings, diagrams, and actions.	2/2

Figure 30. Detail of Individual Student Report (ISR): Science Interim with Scoring Assertions

Marquez, Santos — **SCI Interim Grade 8 - Waves 2018-2019**

Student ID: 991007093 | Student DOB: 2/3/2002 | Enrolled Grade: 8 — Demo District 9998

Date Taken: 10/18/2018 — Demo School 999801

Score: 1/22

Item #	Scoring Assertion	Outcome
1-1	The student correctly identifies the change in amplitude recorded in the simulation, providing some evidence of student ability to observe and summarize how waves change in different media.	✕
1-2	The student correctly identifies that the frequency does not change, providing some evidence of student ability to identify wave properties and how they change in different scenarios.	✕
1-3	The student identifies that frequency does not change, providing some evidence of student ability to identify how properties of the medium affect each wave characteristic.	✕
1-4	The student identified that amplitude increases (or is indeterminate if they did not record it decreasing in the simulation), providing some evidence of student ability to identify how properties of the medium affect each wave characteristic.	✕
1-5	The student indicates that the wavelength cannot be determined simply from density (or that it increases if they found it to decrease in the simulation), providing some evidence of student ability to interpret the data given and make inferences about the affect of media density on each wave characteristic.	✕
1-6	The student correctly calculates and records the amplitude of the wave through salt water (4) providing some evidence of student ability to make observations and use them in a model of sound waves through media with different properties.	✕
1-7	The student correctly calculates the wavelength of salt water (13-15), providing some evidence of student ability to make observations and use them in a model of sound waves through media with different properties.	✕
1-8	The student indicates that the wave speed increases as the density increases based on the observations (or if they recorded the wavelength in salt water as shorter, then decrease, or "cannot tell" if they recorded equal wavelengths), providing some evidence of student ability to use data to identify how wave properties change in each scenario.	✕
1-9	The student correctly calculates and records the the amplitude of the wave through water (5) providing some evidence of student ability to make observations and use them in a model of sound waves through media with different properties.	✕
1-10	The student correctly calculates the wavelength of water (11-13), providing some evidence of student ability to make observations and use them in a model of sound waves through media with different properties.	✕
1-11	The student correctly identifies that the change in wavelength that was recorded in the simulation, providing some evidence of student ability to observe and summarize how waves change in different media.	✕
1-12	The student correctly identifies that the change in wave speed (which goes up with the wavelength recorded in the simulation), providing some evidence of student ability to observe and summarize how waves change in different media.	✕
1-13	The student indicated that the wave speed cannot be determined (or increase, if they found it to decrease in the simulation), providing some evidence of student ability to interpret the data given and make inferences about the effect of media density on each wave characteristic.	✓

Figure 31. Detail of Individual Student Report (ISR): Math Interim (ICA) with Longitudinal Graph

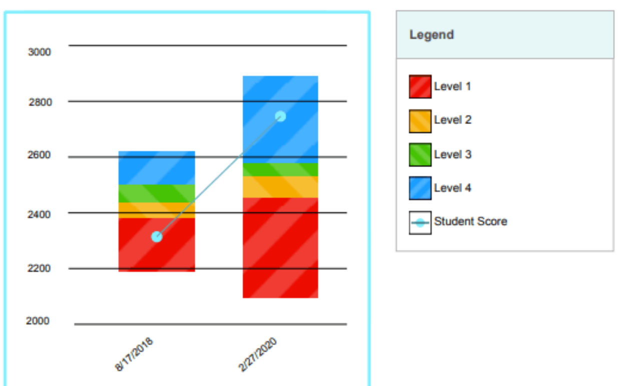
Jackson, Cynthia — **Grade 5 Mathematics - Interim (ICA) - Combined 2019-2020**

Student Name: Jackson, Cynthia | Student ID: 99996727 | Enrolled Grade: 5 — TestDISTRICT12

Date Taken: 2/27/2020 — TestSCHOOL9

Achievement Level: Level 4 **Scale Score:** 2746±58

Trend Report



Date	Test Reason	Test Label	Scale Score	Performance Level
8/17/2018 12:00:00 AM	Unassigned	Grade 3 MATH -COMBINED	2314±6	Level 1
2/27/2020 12:00:00 AM	Unassigned	Grade 5 Mathematics - Interim (ICA) - Combined	2746±58	Level 4

How to Generate and Export Individual Student Reports (ISR)

To generate and export an Individual Student Report (ISR) that details a student's performance on a test opportunity, use the Student Results Generator. You can select any combination of a test reason, assessments within a subject, and students in order to generate either a single ISR or multiple ISRs at once.

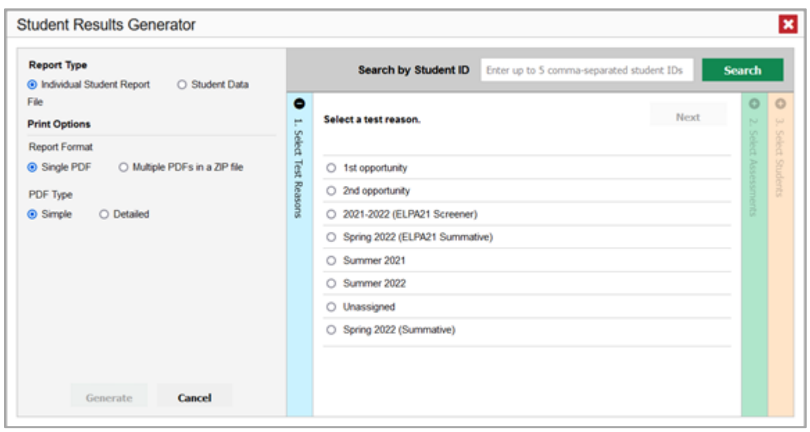
You may want to use the Student Results Generator to simultaneously print large numbers of ISRs.

ISRs can be generated from almost any Reporting page.

1. Click the **Download Student Results** button in the **Features & Tools** menu. The ***Student Results Generator*** window opens (Figure 32).

   Depending what page you open the Student Results Generator from, the options available to you may be prepopulated or preselected; for example, the Student Portfolio Report prepopulates a single student, and the Student Performance on Test report preselects a single test opportunity. (The filters applied to the page have no effect, however.) You can change the selections.

Figure 32. Student Results Generator Window as Opened from Performance on Tests Report



2. In the panel on the left, select **Individual Student Report**. Always do this before you make other selections. Switching between the **Individual Student Report** and **Student Data File** options may revert some selections.

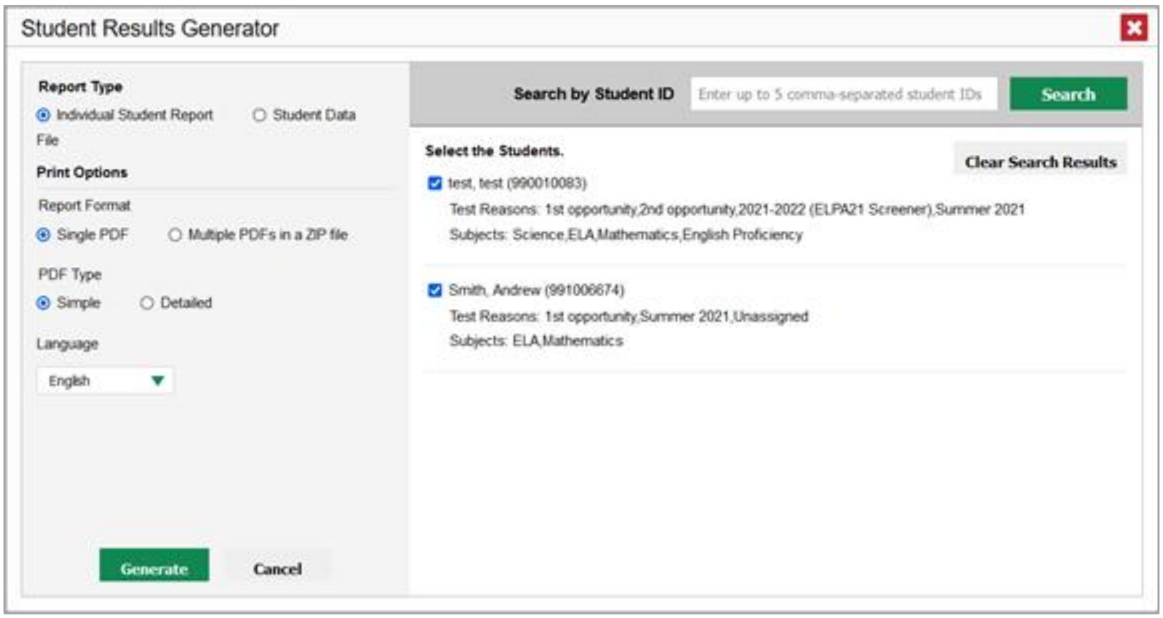
3. If you're generating multiple ISRs, then under Report Format, choose either a single PDF for all the ISRs, or a ZIP file containing a separate PDF for each one. If you select **Single PDF**, the Student Results Generator may nonetheless create a ZIP file of multiple PDFs depending on the number of schools, grades, and opportunities included.

4. Under PDF Type, select either a simple or detailed PDF.

5. If the test opportunity options are not preselected, or if you want to change them, there are two ways to make selections:

- Search for students. In the search field at the upper-right corner, enter up to 5 comma-separated student IDs and click **Search**. The resulting list of students and all the tests they've taken will replace any previous selections, as in Figure 33. To deselect and clear results, click **Clear Search Results**.

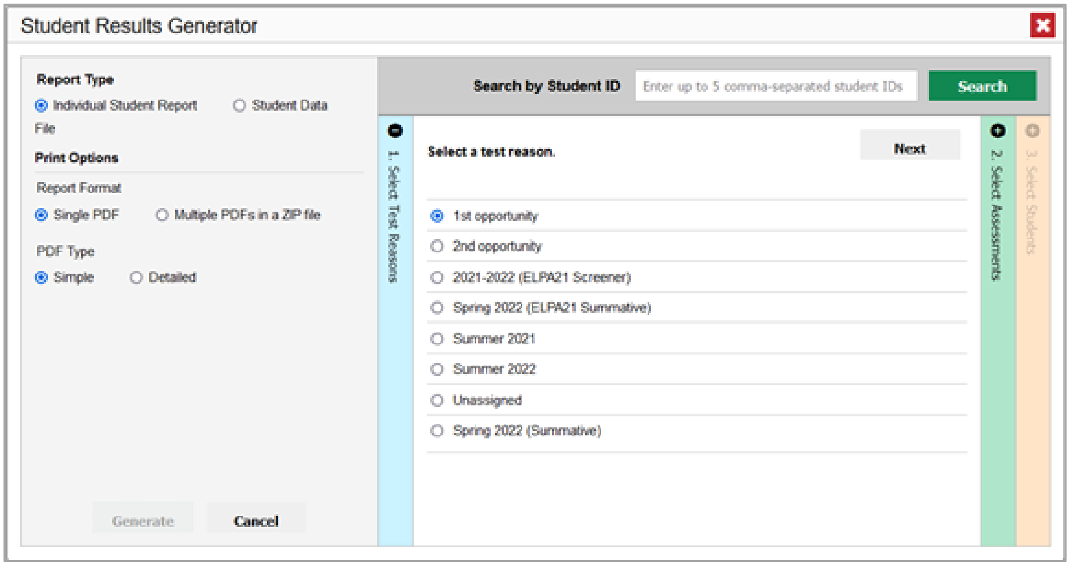
Figure 33. Student Results Generator Window: Student Search Results



- Use the three accordion sections. (If a student is prepopulated by the Student Portfolio Report, first remove the student by clicking **Clear Search Results**.) Starting from the left, click the section bars to expand them or use the **Next** and **Previous** buttons to navigate. Within each section you must make selections using the radio buttons and checkboxes:

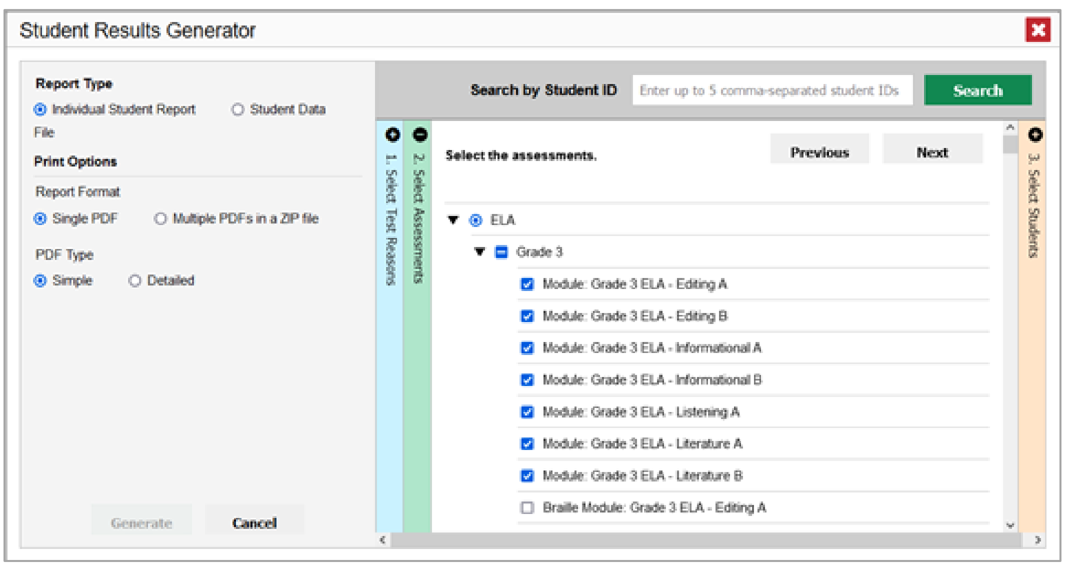
i. In the **Select Test Reason** accordion section (Figure 34), choose a test reason. Test reasons are either test windows or categories for tests.

Figure 34. Student Results Generator Window: Select Test Reason Section



ii. In the **Select Assessments** section (Figure 35), choose any number of tests or grade levels within a single subject.

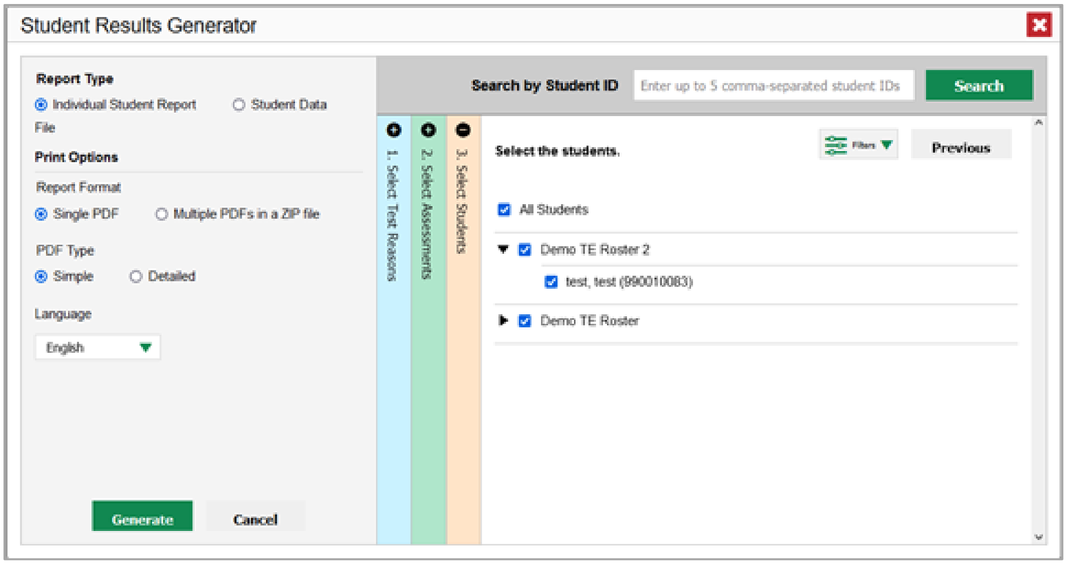
Figure 35. Student Results Generator Window: Select Assessments Section



iii. In the **Select Students** section (Figure 36), choose students from the expandable school and/or class (roster) options. If you're a district-level user, you may choose students in up to three schools.

- Sometimes the list of students is truncated. You can display the entire list by clicking **Click to Load More**.
- Note that marking the checkbox for a student in one class (roster) or school also marks it anywhere else the student appears, and the same goes for clearing the checkbox.

Figure 36. Teacher View: Student Results Generator Window: Select Students Section



6. *Optional*: If you used the accordion sections to make your selections, then to set a range of processing dates for which to generate results, use the filter menu as follows:

a. Open the **Filters** menu . The menu displays two date fields, as in Figure 37.

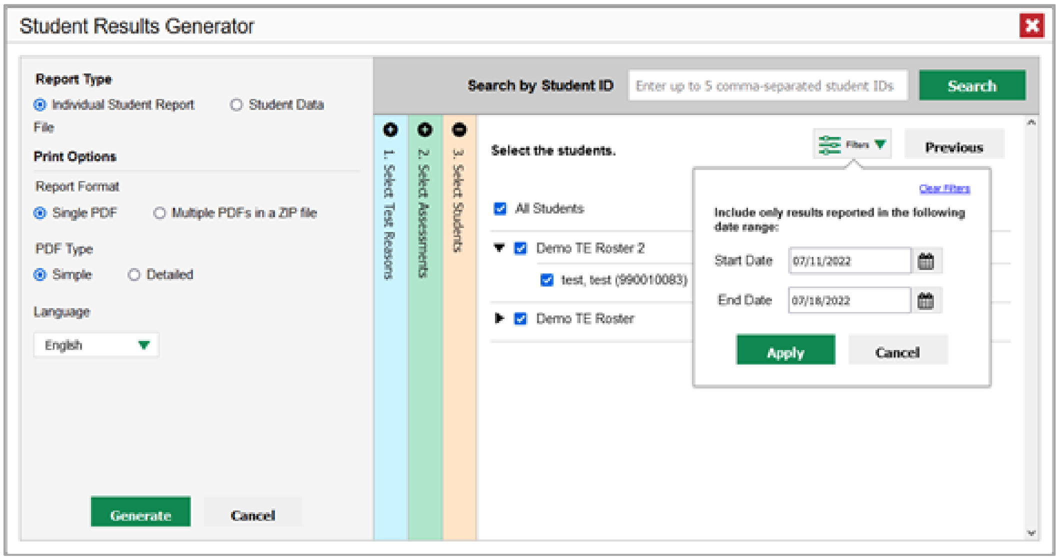
b. Use the calendar tools to select dates, or enter them in the format `mm/dd/yyyy`.

c. Click **Apply**.

d. *Optional*: To revert to including results for all available dates, reopen the filter menu, click **Clear Filters**, then click **Apply**.

Note that processing date is not always the same as the date a test was taken.

Figure 37. Teacher View: Student Results Generator Window: Select Students Section with Filters Menu Open



7. Click **Generate**. Once ISR generation is finished, the Inbox contains the new ISR(s) available for download.

Note that if a student took a test multiple times with different test reasons, an ISR will be generated for each test opportunity. If a student took a test multiple times with the same test reason, only the most recent test opportunity will be included. You can create an ISR for another test opportunity by navigating directly to the report for that opportunity. Past test opportunities are marked with numbers (1) in reports, starting with the earliest.

How to Generate and Export Student Data Files

This section discusses student data files, which are useful for analysis.

To generate and export student data files, use the Student Results Generator. You can select any combination of a test reason, assessments, and students in order to generate and export the files.

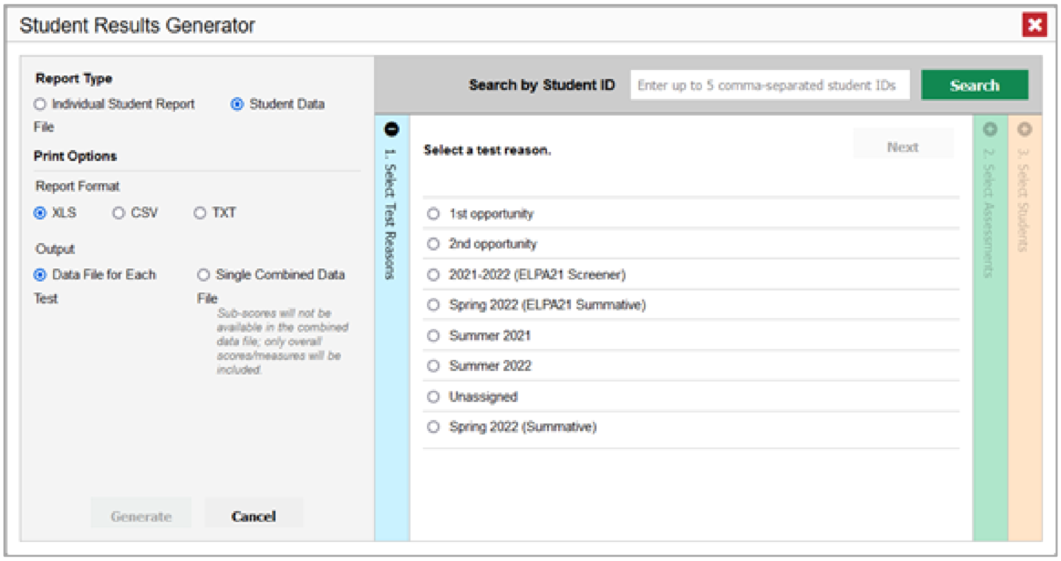
You can generate student data files from almost any report page.

1. Click the **Download Student Results** button in the **Features & Tools** menu. The ***Student Results Generator*** window opens.

   Depending what page you open the Student Results Generator from, the options available to you may be prepopulated or preselected; for example, the Student Portfolio Report prepopulates a single student, and the Student Performance on Test report preselects a single test opportunity. (The filters applied to the page have no effect, however.) You can change the selections.

2. In the panel on the left, select **Student Data File**, as in Figure 38. Always do this before you make other selections. Switching between the **Individual Student Report** and **Student Data File** options may revert some selections.

Figure 38. Student Results Generator Window



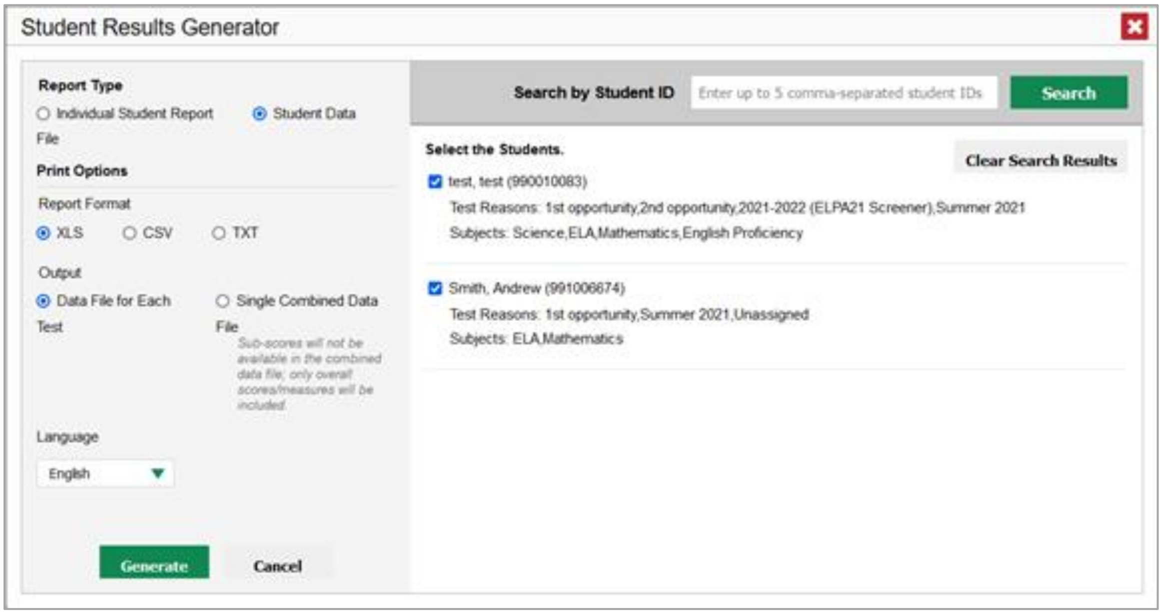
3. Under Report Format, select **XLS** (Excel .xlsx), **CSV** (comma-separated values), or **TXT** (tab-delimited text).

4. Under Output, select either **Data File for Each Test** or **Single Combined Data File**. Note that a single combined file does not include reporting categories. Large files may be split by school.

5. If the test opportunity options are not preselected, or if you want to change them, there are two ways to make selections:

   - Search for students. In the search field at the upper-right corner, enter up to 5 comma-separated student IDs and click **Search**. The resulting list of students and all the tests they've taken will replace any previous selections, as in Figure 39. To deselect and clear results, click **Clear Search Results**.

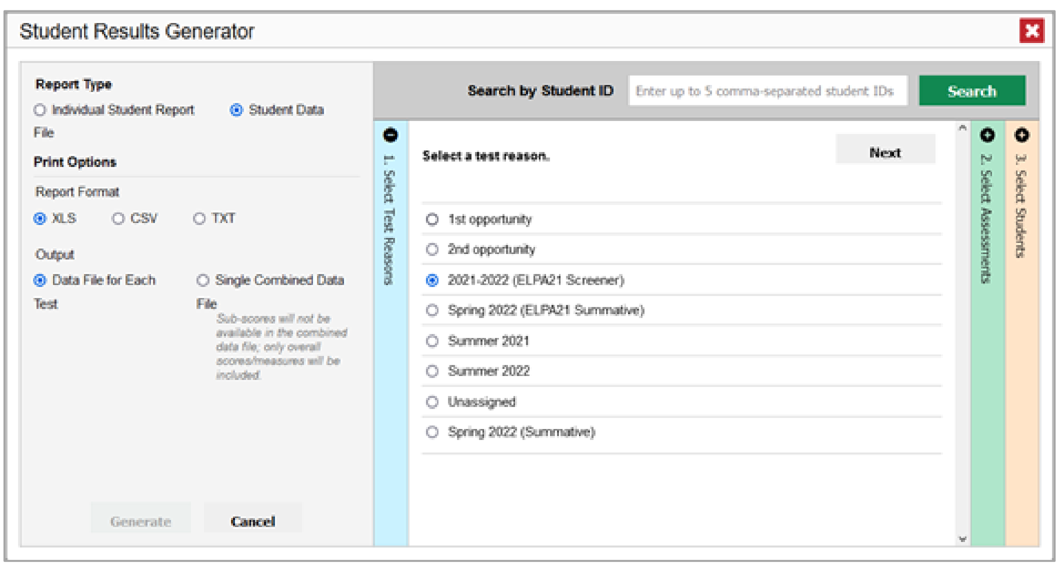
Figure 39. Student Results Generator Window: Student Search Results



- Use the three accordion sections. (If a student is prepopulated by the Student Portfolio Report, first remove the student by clicking **Clear Search Results**.) Starting from the left, click the section bars to expand them or use the **Next** and **Previous** buttons to navigate. Within each section you must make selections using the radio buttons and checkboxes:

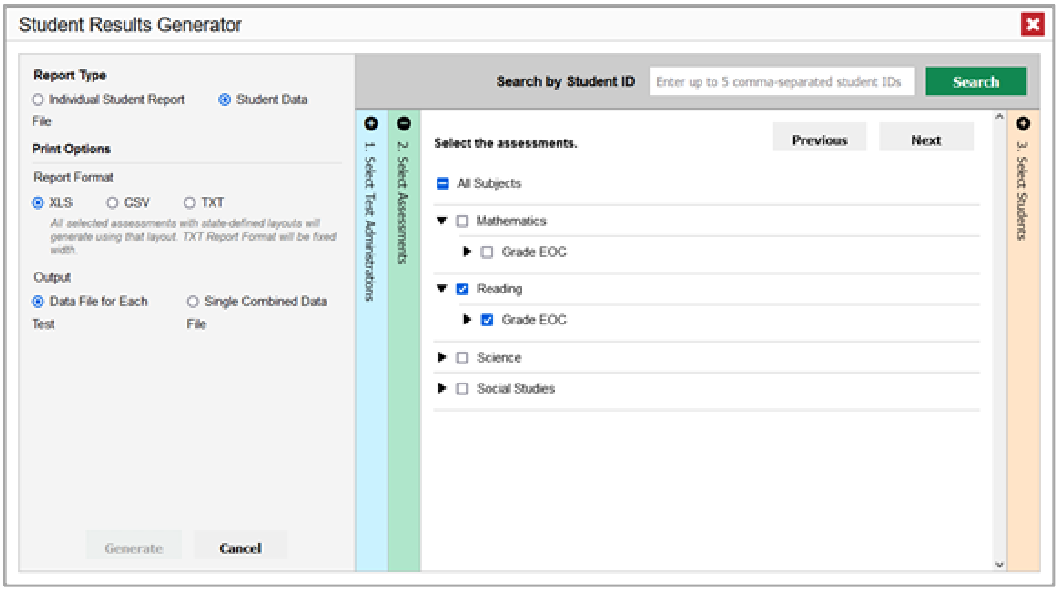
  i. In the **Select Test Reason** section (Figure 40), choose a test reason. Test reasons are either test windows or categories for tests.

Figure 40. Student Results Generator Window: Select Test Reason Section



  ii. The **Select Assessments** section (Figure 41) groups tests by subject and grade. Mark the checkboxes beside the tests or groups of tests you want to include in the report, or mark **All Subjects**.

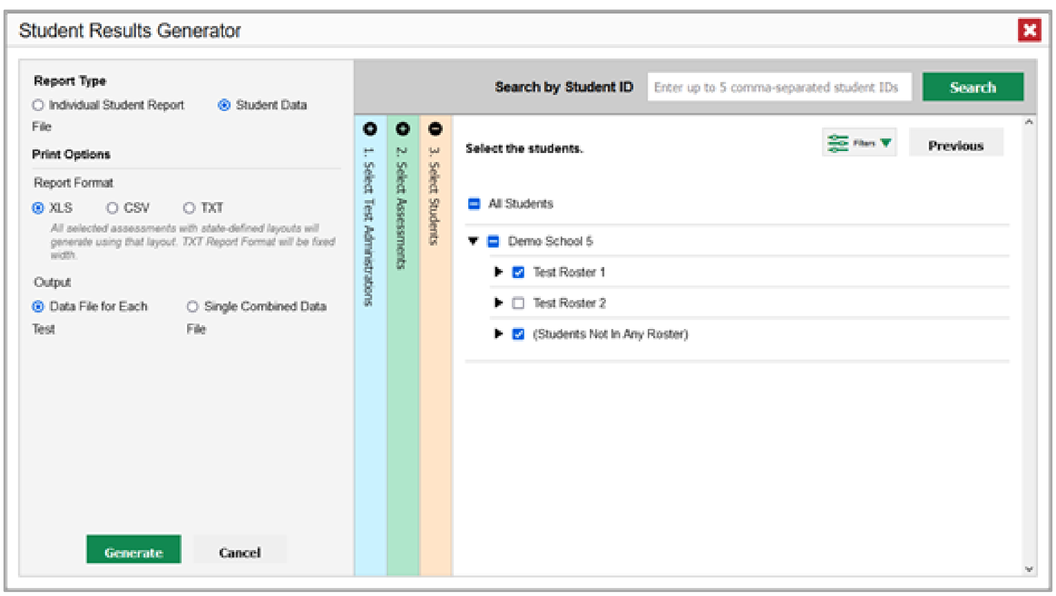
Figure 41. Student Results Generator Window: Select Assessments Section



iii. In the **Select Students** section (Figure 42), select any number of students from the expandable school and/or class (roster) options.

- Sometimes a list of students is truncated. You can display the entire list by clicking **Click to Load More**.
- Marking the checkbox for a student in one class (roster) or school also marks it anywhere else the student appears, and the same goes for clearing the checkbox.

Figure 42. District-Level User View: Student Results Generator Window: Select Students Section



7. *Optional*: If you used the accordion sections to make your selections, then to set a range of processing dates for which to generate results, use the filter menu as follows:

a. Open the **Filters** menu (see Figure 43). The menu displays two date fields.

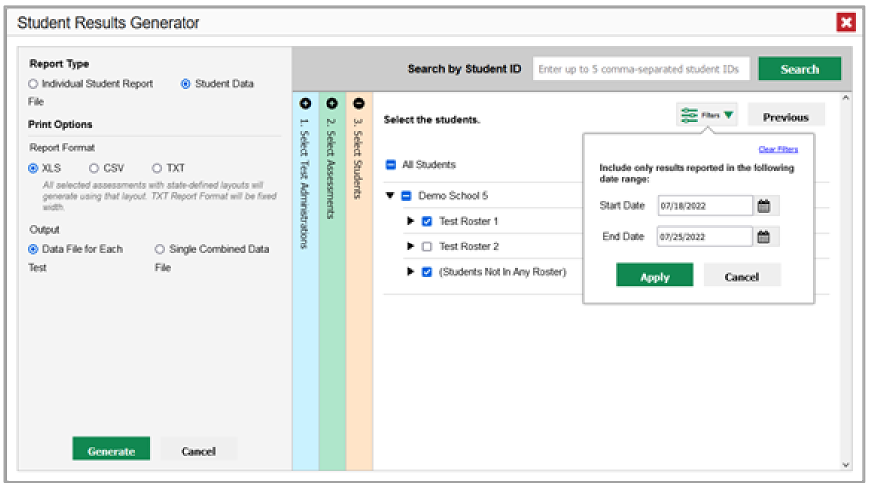
b. Use the calendar tools to select dates, or enter them in the format `mm/dd/yyyy`.

c. Click **Apply**.

d. *Optional*: To revert to including results for all available dates, reopen the filter menu, click **Clear Filters**, then click **Apply**.

Note that processing date is not always the same as the date a test was taken.

Figure 43. District-Level User View: Student Results Generator Window: Select Students Section with Filters Menu Open



8. Click **Generate**. Once data file generation is finished, the Inbox contains the new student data file(s) available for download.

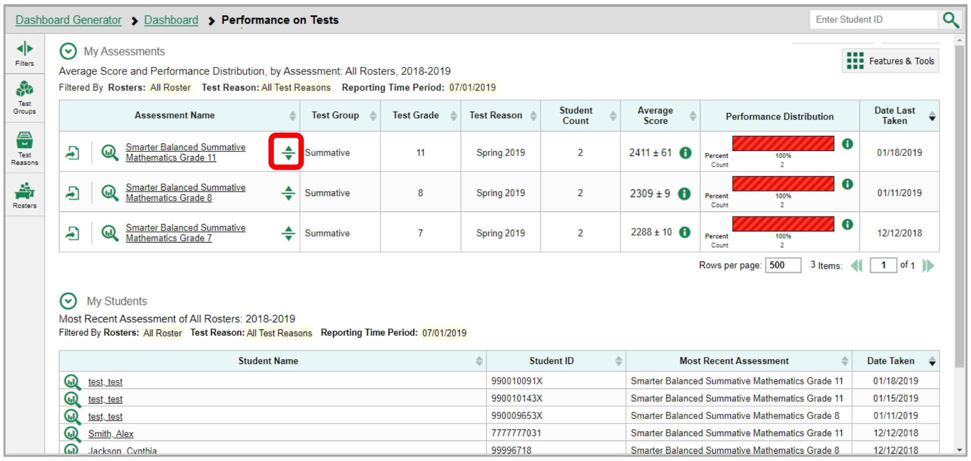
Note that if a student took a test multiple times, the files will include each test opportunity.

How to Compare Students' Data with Data for Your State, District, School, and/or Total Students

In the Performance on Tests report and the Student Portfolio Report, you can access performance data for your state, district, school, and/or total students.

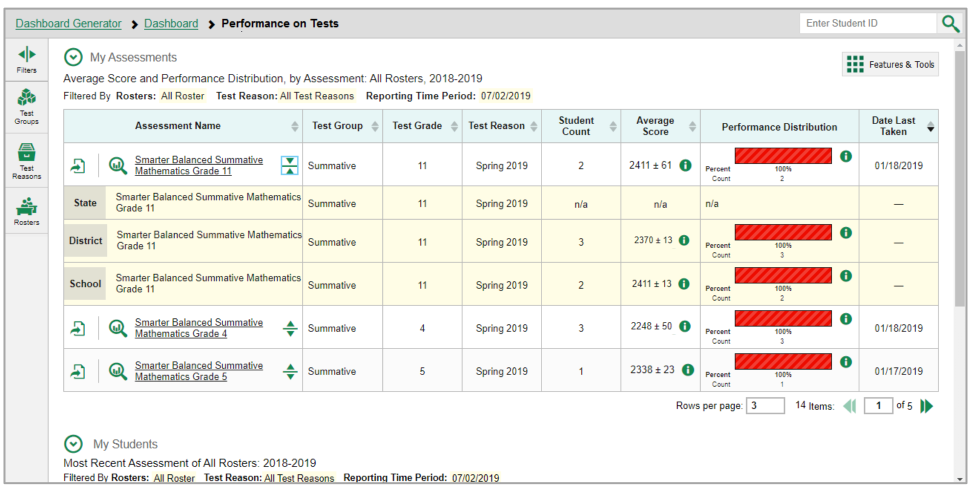
In the Performance on Tests report (see Figure 44), click to the right of a test name.

Figure 44. Teacher View: Performance on Tests Report



Rows containing data for the state, district, and/or school appear below, as in Figure 45.

Figure 45. Teacher View: Performance on Tests Report with Expanded Comparison Rows



To hide the comparison rows, click to the right of the test name.

The comparison feature also appears on the Student Portfolio Report page, to generate comparisons for a student's performance on any test with that of your state, district, school, and/or total students.

How to Set Up Reports for Summatives and Interims to Suit Your Needs

You can set up your reports so it's easier to access the data that are most important to you. For example, if you're a teacher, you may want to hide certain tests in subjects you don't teach, or you may want to narrow down your reports to a single roster.

This section explains how to make several different adjustments to reports: showing only the tests you're interested in; showing only the classes (rosters) you're interested in; showing only the schools you're interested in; and viewing data from a previous point in time.

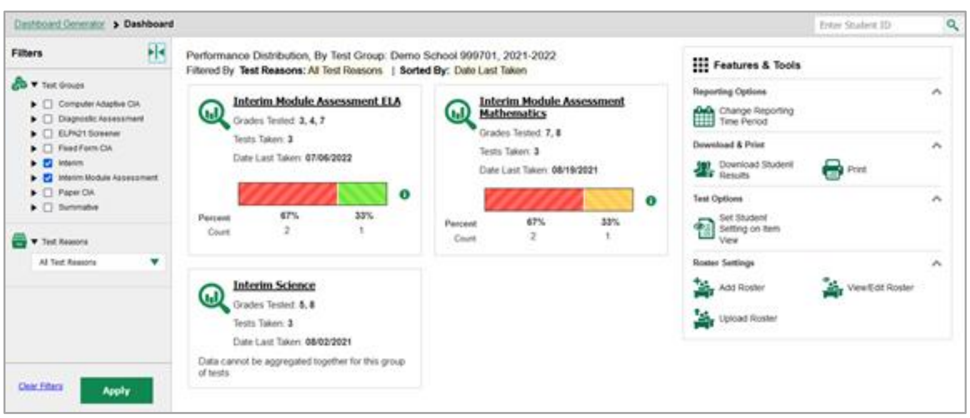
How to Filter Tests to Display

You can filter the tests you want to view in reports. You may want to do this, for example, if you are an ELA teacher and you do not want to see your students' math scores. By default, the data for those mathematics assessments appear in your reports.

Filtering tests to display begins on the Dashboard Generator page. This is where you can select the test groups you want to view on the dashboard and, as an option, set those selections as defaults. You can also temporarily filter the tests that appear in the reports you are currently viewing, as described below.

1. On the left side of the dashboard or the Performance on Tests report, click either the **Filters** panel expand button or the **Test Groups** button. The **Filters** panel expands (see Figure 46).
2. Mark as many selections as you like in the **Test Groups** section of the filters panel (see Figure 46). Tests are organized by test type, subject, and grade.

Figure 46. Teacher View: Dashboard with Expanded Filters Panel



3. Click **Apply**. The report updates to show only data for those tests.
4. *Optional*: To revert all filters to their defaults, open the filters panel again and click **Clear Filters**. Click **Apply**. Filters also revert when you log out, switch user roles, or switch systems.

For Teachers and School-Level Users: How to Filter Classes (Rosters) to Display

Rosters usually represent classes, but can represent any group that's meaningful to users, such as students who have taken honors courses. Each roster is associated with a teacher. To learn how to create and modify rosters in the Reporting System, see Class (Roster) Management.

In the Performance on Tests report, teachers and school-level users can filter by a particular roster. When you filter, you eliminate students not in the selected class from the data you're viewing.

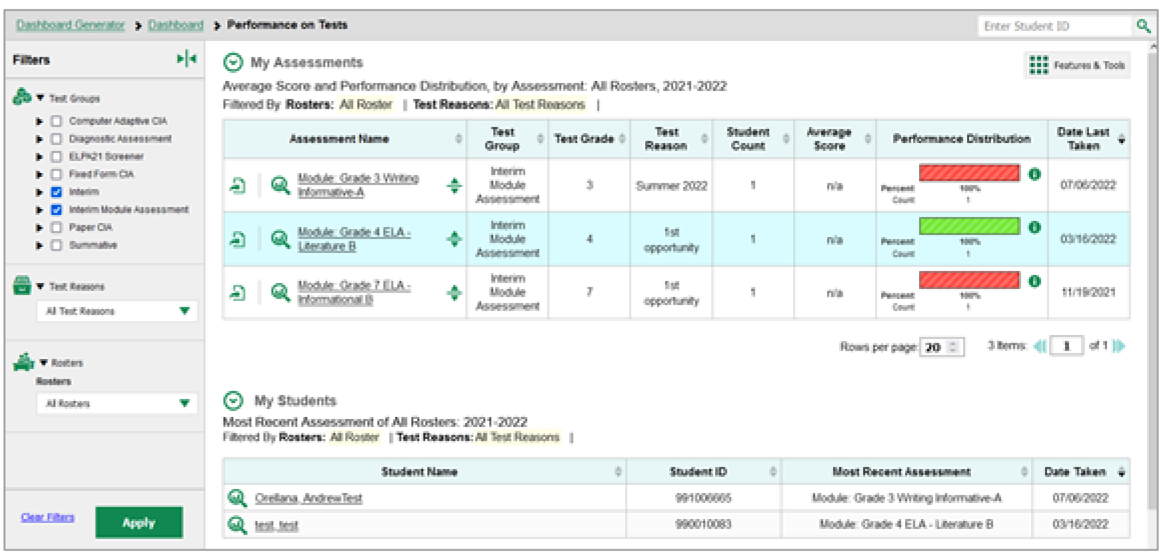
Filtering by roster makes it easy to focus on a particular class's performance. And by switching filters, you can easily compare one class with another. If you don't filter by roster, the reports default to showing data for all classes. You may find data for a single class easier to understand.

1. On the left side of the Performance on Tests report, click either the **Filters** panel expand button or the **Rosters** button. The **Filters** panel expands (see Figure 47).

2. Make a selection from the drop-down list in the **Rosters** section.

   - If you're a school-level user, you must first select a teacher from the drop-down list, and then select a particular class (roster) from the second drop-down list that appears. By default the first class listed is selected.

Figure 47. Teacher View: Performance on Tests Report with Expanded Filters Panel



3. Click **Apply**. The report updates to show only data for that class (roster).

4. *Optional*: To revert all filters to their defaults, open the filters panel again and click **Clear Filters**. Click **Apply**. Filters also revert when you log out, switch user roles, or switch systems.

All the reports accessible from this page will be filtered the same way.

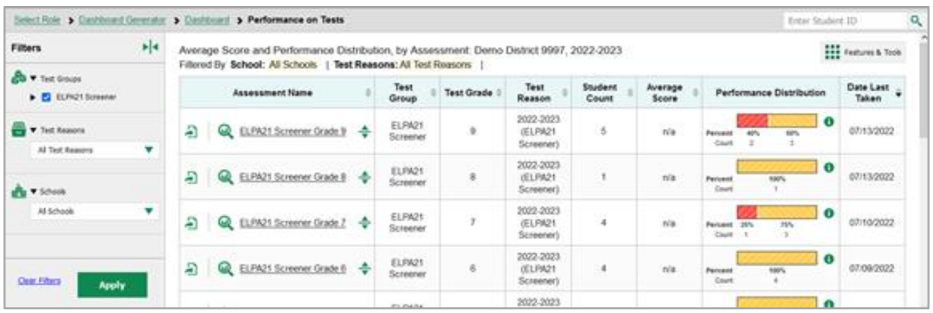
The row of filter details below each table header shows the classes (rosters) you're viewing.

For District-Level Users: How to Filter Schools to Display

Filtering the Performance on Tests report by school makes it easy to focus on a particular school's performance. And by switching filters, you can easily compare it with another school. If you don't filter by school, the Performance on Tests report defaults to showing data for all schools. You may find data for a single school easier to understand.

1. On the left side of the Performance on Tests report, click either the expand button or the **Schools** button. The **Filters** panel expands (see Figure 48).
2. Make a selection from the drop-down list in the **Schools** section (see Figure 48).

Figure 48. District-Level User View: Performance on Tests Report with Expanded Filters Panel



3. Click **Apply**. The report updates to show only data for that school.
4. *Optional*: To revert all filters to their defaults, open the filters panel again and click **Clear Filters**. Click **Apply**. Filters also revert when you log out, switch user roles, or switch systems.

If you click the name of a test (or beside it) when you've filtered by a single school, the link will take you to the School Performance on Test report and not to the District Performance on Test report.

The row of filter details below the table header shows the schools you're viewing.

How to View Data from a Previous Point in Time

Changing the reporting time period allows you to view test results from a previous point in time. There are two time period settings: you can select a school year for which to view tests, and you can enter a date for which to view students.

- When you set a school year for which to view tests, the reports show data for test opportunities completed *in the selected school year*.
- When you set a date for which to view students, the reports show data only for the students who were associated with you *as of the selected date*. Students' enrollment and demographic

information is all given as of the selected date as well. You can use this setting to view data for students who have left or recently entered your classes (rosters), school, or district.

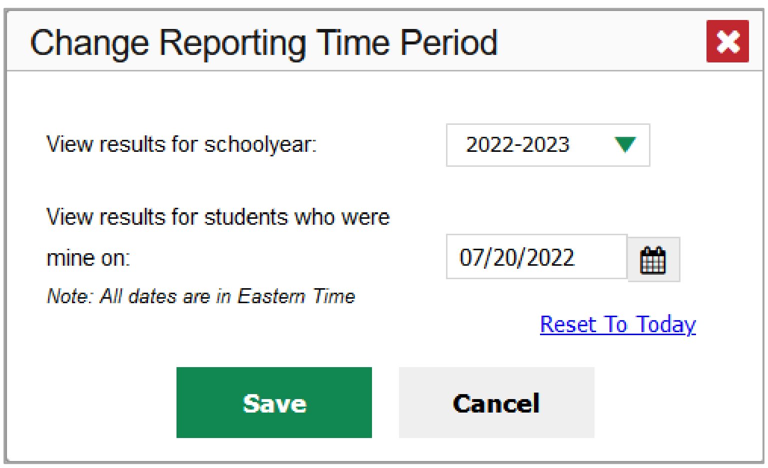
If you don't change the reporting time period, or if you reset it to the default, all the reports show test opportunities only for the current school year (except Longitudinal Reports and Student Portfolio Reports, which always retain the ability to look back to previous years), with current student data.

Some examples of how you can use this feature:

- You may want to view the past performance of your current students, including new transfer students. In that case, set a school year in the past and keep the date set to today.
- You may want to view the performance of your former students in order to compare them with that of your current students. In that case, set the date to a time when your former students belonged to you and had started testing, and set the school year to the same time. Then switch back to the present to compare.

1. From the **Features & Tools** menu, select **Change Reporting Time Period**. (If you're viewing the Dashboard Generator page, click **Change the reporting time period.** The dashboard also offers this link when no assessments are available to display.) The ***Change Reporting Time Period*** window appears (see Figure 49).

Figure 49. Change Reporting Time Period Window



2. From the school year drop-down list, select a school year (see Figure 49). This is the year for which you will view test results.

3. In the *View results for students who were mine on* field, use the calendar tool to select a date, or enter it in the format `mm/dd/yyyy`. You will be viewing all the students who were associated with you on that date, and only those students.

    - To view your current students' past performance, keep the date set to today.
    - To view the performance of your former students, set the date to a day when those students were associated with you and had started testing.

4. Click **Save**. All reports are now filtered to show only data for the selected school year and date. The selected date displays in the filter details below the report headings. All other filters are cleared.

5. *Optional*: To go back to viewing the latest data, open the ***Change Reporting Time Period*** window again, click **Reset To Today** in the lower-right corner, then click **Save**. The date resets and all filters are cleared. The reporting time period also resets when you log out, but persists when you switch roles.

How to Export and Print Data on Summatives

You can export or print any data you see in the Reporting System. Some can be exported directly from the Performance on Tests report. You may want to export or print to save a snapshot of data to consult later, or to share data. Different options will be available depending on the report you are viewing.

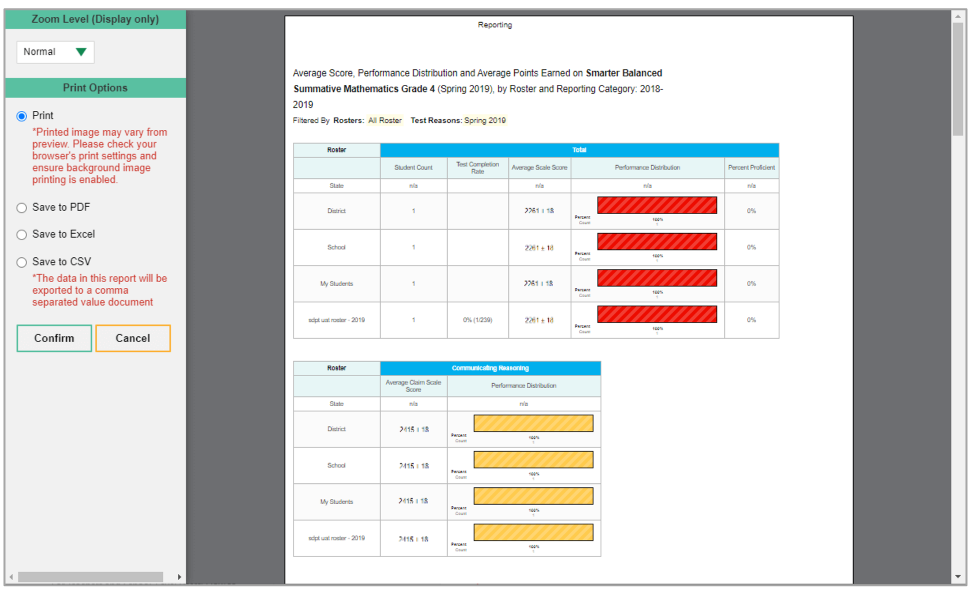
How to Export or Print a Report You're Viewing

1. Select the print button from the **Features & Tools menu** or, in a Longitudinal Report window, from the upper-right corner. If there are multiple report tables on the page, multiple print buttons appear.

   A print preview page opens (see Figure 50).

   - To zoom in on the print preview, use the drop-down list under the *Zoom Level (Display only)* section. This setting affects the preview only.

Figure 50. Print Preview Page



2. Do one of the following under the *Print Options* section:

   - To print the report, select the **Print** radio button.
   - To download a PDF version of the report, select **Save to PDF**. Then select an option from the **Page Layout** drop-down list that appears.
   - To download a Microsoft Excel (.xlsx) version, select **Save to Excel**.
   - To download a comma-separated value (CSV) version of the report, select **Save to CSV**.

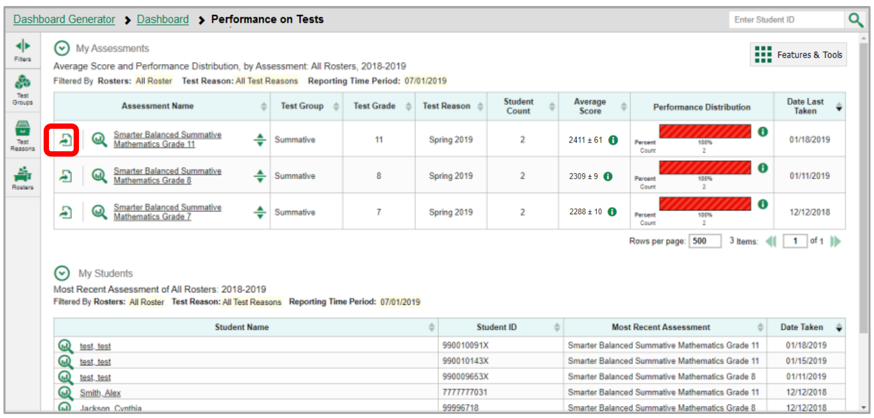
3. Click **Confirm**.

If you saved the report as PDF, Excel, or CSV, the ***Secure Inbox*** window appears, displaying the generated report.

How to Export an Assessment Report Directly from the Performance on Tests Report

1. Click the export button to the left of the name of the assessment whose report you wish to export (see Figure 51).

Figure 51. Teacher View: Performance on Tests Report

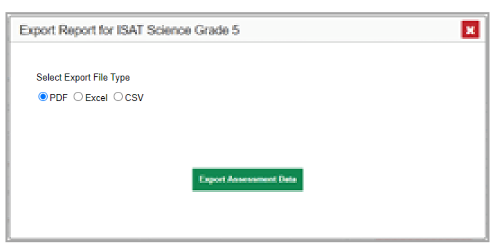


The ***Export Report*** window opens. The options in this window vary according to your user role (see Figure 52 and Figure 53).

2. If necessary, select which report to export for the assessment.

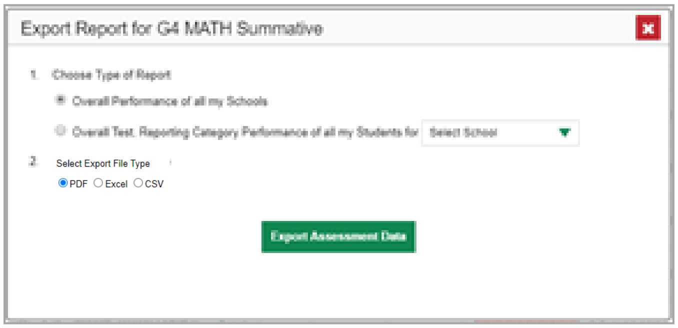
- **Teachers and school-level users**: The exported report will contain test results for all your students.

Figure 52. Teacher View: Export Report Window



- **District-level users**: Select which report to export for the assessment (see Figure 53).
  - To export the district test results for the assessment, mark the **Overall Performance of all my Schools** radio button.
  - To export school test results, mark the **Overall Test & Reporting Category Performance of all my Students for [School Name]** radio button, then select a school from the drop-down list.

Figure 53. District-Level User View: Export Report Window



5. Do one of the following:
   - To export the report in PDF format, mark the **PDF** radio button.
   - To export the report in .xlsx format, mark the **Excel** radio button.
   - To export the report in comma-separated values (CSV) format, mark the **CSV** radio button.
6. Click **Export Assessment Data**. A confirmation window appears.
7. Click **Yes** to export or **No** to return to the ***Export Report*** window. When you've exported a file, the ***Secure Inbox*** window appears with the generated file available for download.

More About How to Work with Interims

This section explains some Reporting System features and functions that are specific to interim assessment reports. These features cannot be used with summative assessment reports.

How to Access Item-Level Data on Interims

Unlike summatives, interim and benchmark assessments contain non-secure, non-public items. Reports for individual interim and benchmark tests include the following:

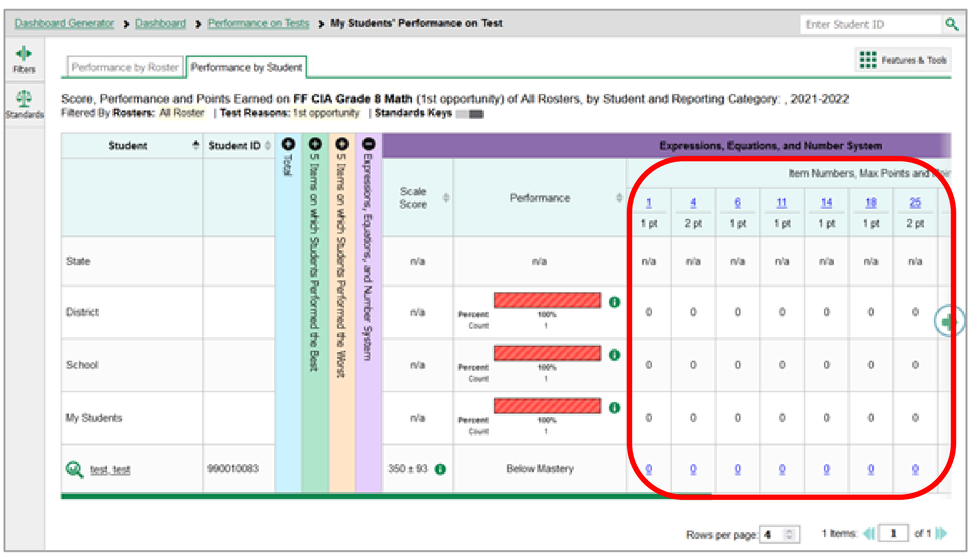
- Item-level data.
- Access to the items themselves.
- Access to student responses to the items.

Test results for adaptive assessments include item-level data only on the individual student level.

How to View Item Scores

To expand sections containing item data, click the vertical section bars as in Figure 54.

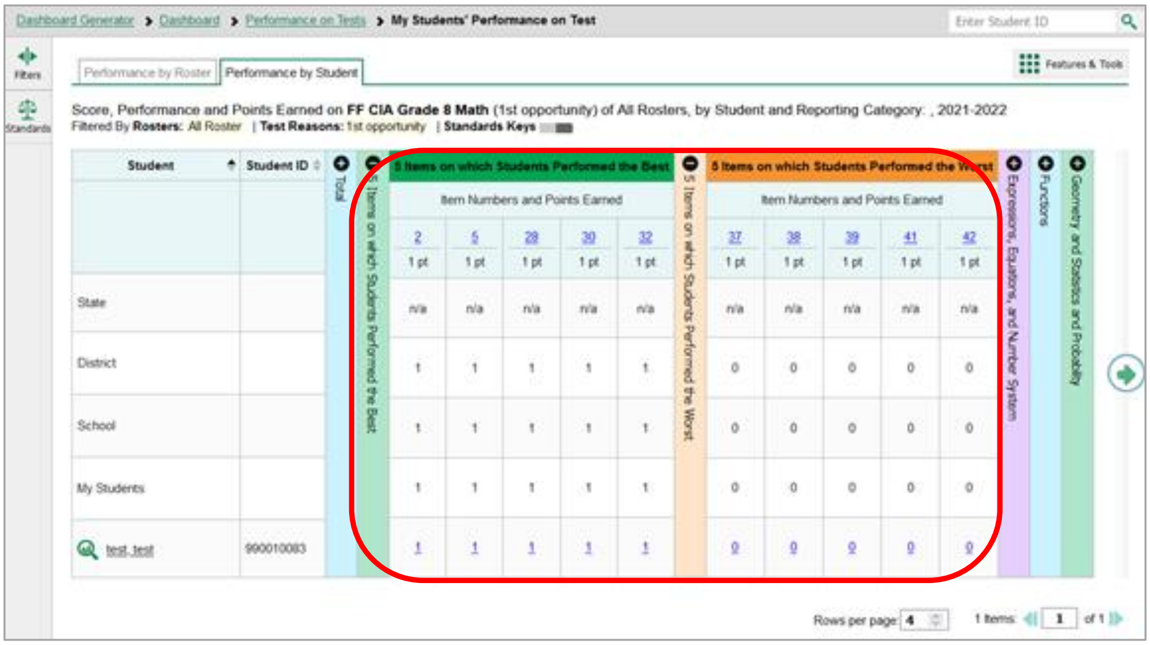
Figure 54. My Students' Performance on Test Report: Performance by Student Tab with Expanded Reporting Category Section



How to Find Out Which Items Students Performed on the Best or Struggled with the Most

Look in the sections **5 Items on Which Students Performed the Best** and **5 Items on Which Students Performed the Worst** (see Figure 55). You can click the vertical section bars to expand them, just like other sections.

Figure 55. My Students' Performance on Test Report: Performance by Student Tab with Expanded 5 Items on Which Students Performed the Best and Worst Sections

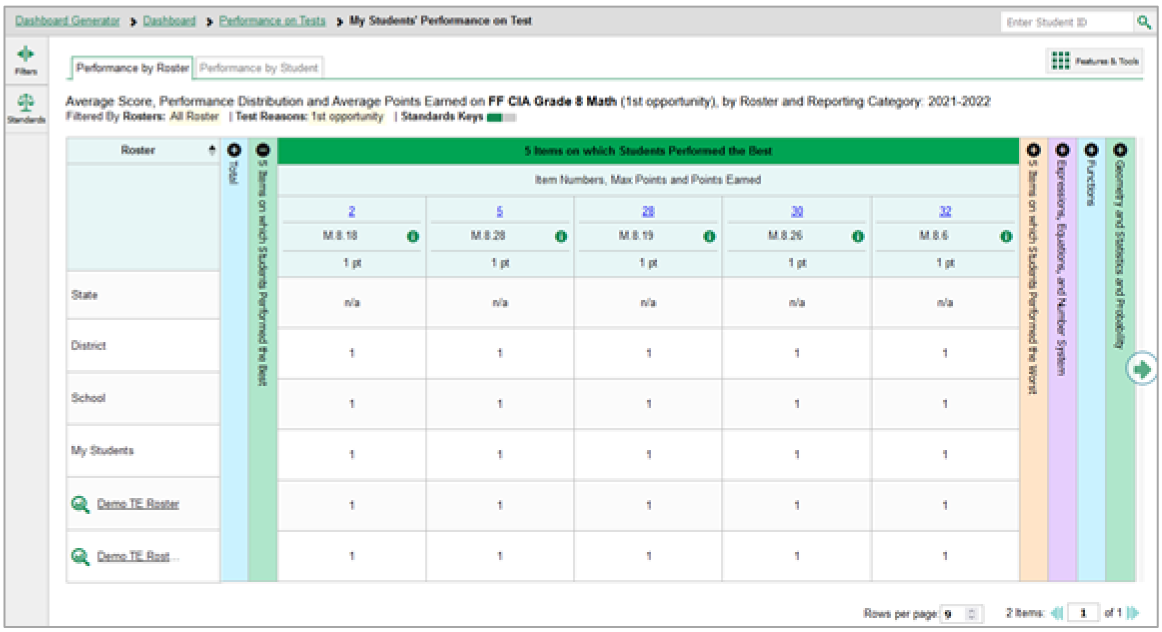


How to View Standards for Each Item

In a report displaying item-level data, you can view the standard or standards to which each item is aligned. This allows you to determine at a glance what the item measures.

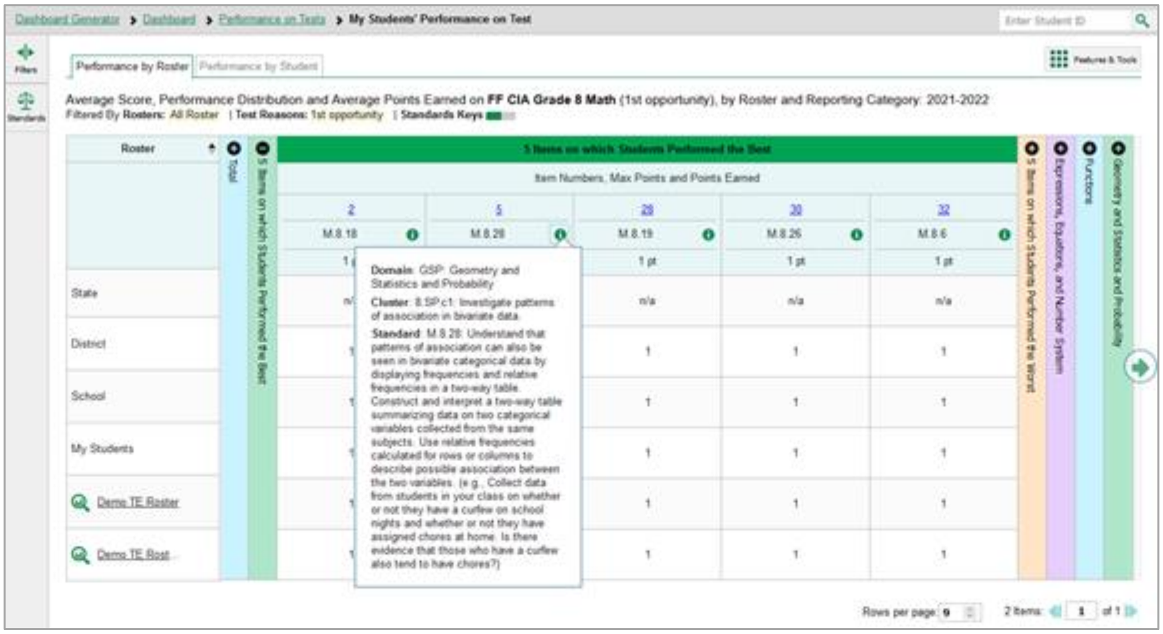
To show and hide item standards, click the **Standards Keys** toggle in the row of filter details below the report table heading. Under each item number appears a standard key or list of standard keys (see Figure 56). Note that this toggle does not affect printouts or exports, which always include the standard keys when they include item-level data.

Figure 56. My Students' Performance on Test Report with Expanded 5 Items on Which Students Performed the Best Section



Click the more information buttons beside the standard keys to view legends displaying the full text of each cluster (category of standards) and each standard, as in Figure 57. This full text is not included in printouts or exports.

Figure 57. My Students' Performance on Test Report with Expanded Reporting Category Section and Expanded Legend



How to View an Item

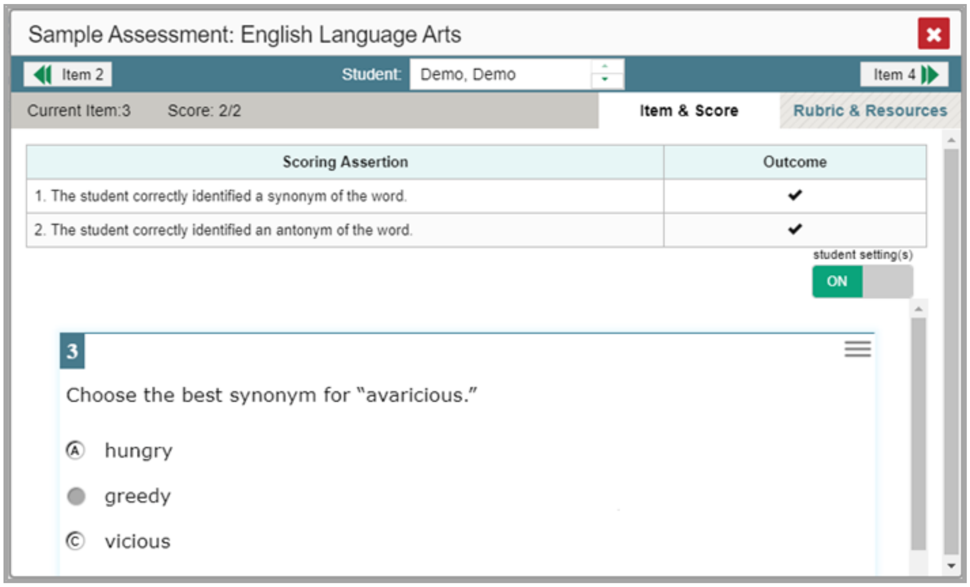
You can view the actual items themselves, along with student responses to those items.

Do either of the following:

- To view the item in a blank state, click the item number in the first row of the report table.
- To view the student's response to the item, find that student's name in the Student column on the left. Then click the score the student obtained on that item.

The ***Item View*** window appears (see Figure 58). It contains an **Item & Score** tab and a **Rubric & Resources** tab. A banner at the top of the window displays the item's number, score (when the item includes the student's response), and confidence level (when a machine-suggested score has a low confidence level). The **Item & Score** tab (see Figure 58) shows the item and may include a particular student's response.

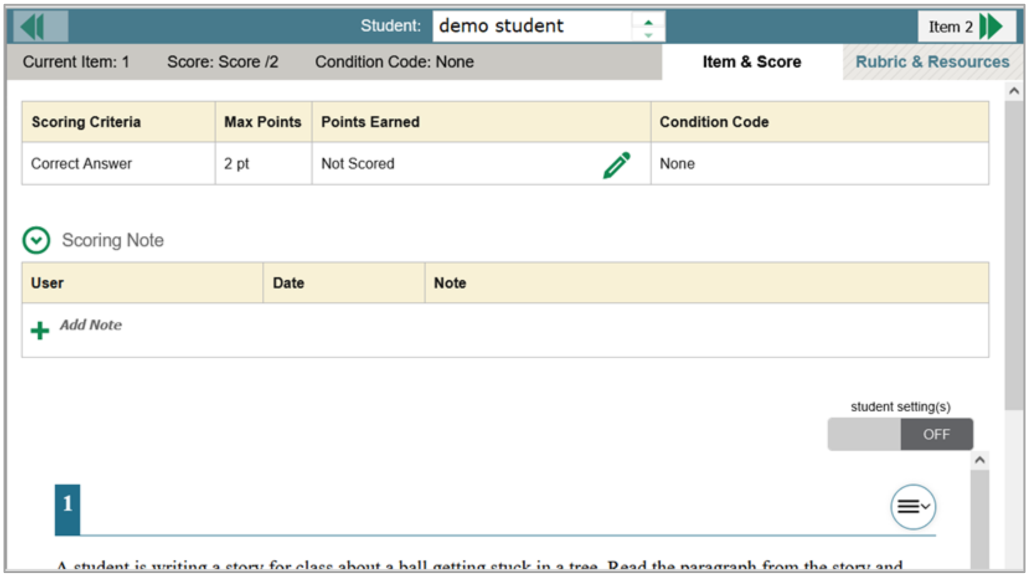
Figure 58. Item View Window: Item & Score Tab with Student Response



The **Item & Score** tab may include the following sections.

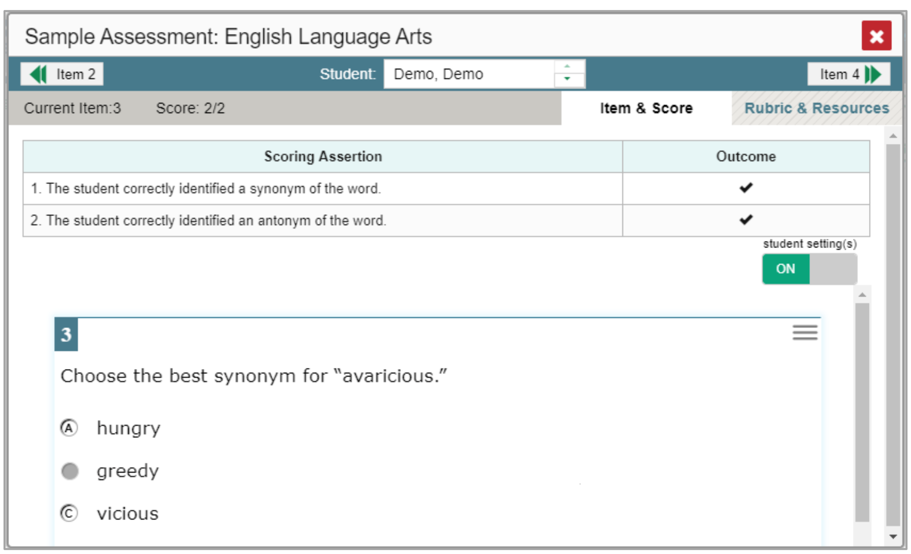
- **Scoring Criteria:** When you're viewing a student's response and the item has scoring criteria, the Scoring Criteria table (see Figure 59) lists the name, maximum points, points earned, and condition codes for each scoring criterion. This table also allows you to modify scores for items with editable scores. Note that for some items, a second Scoring Criteria table appears, displaying transformed scores.

Figure 59. Item View Window: Item & Score Tab with Student Response and Scoring Criteria Table



- **Scoring Assertion:** Each scoring assertion contains both a statement that provides information about what the student did in their response, and the content knowledge, skill, or ability that is evidenced by their response. When you're viewing a student's response and the item has scoring assertions, the Scoring Assertion table appears, listing each assertion and outcome (see Figure 60).

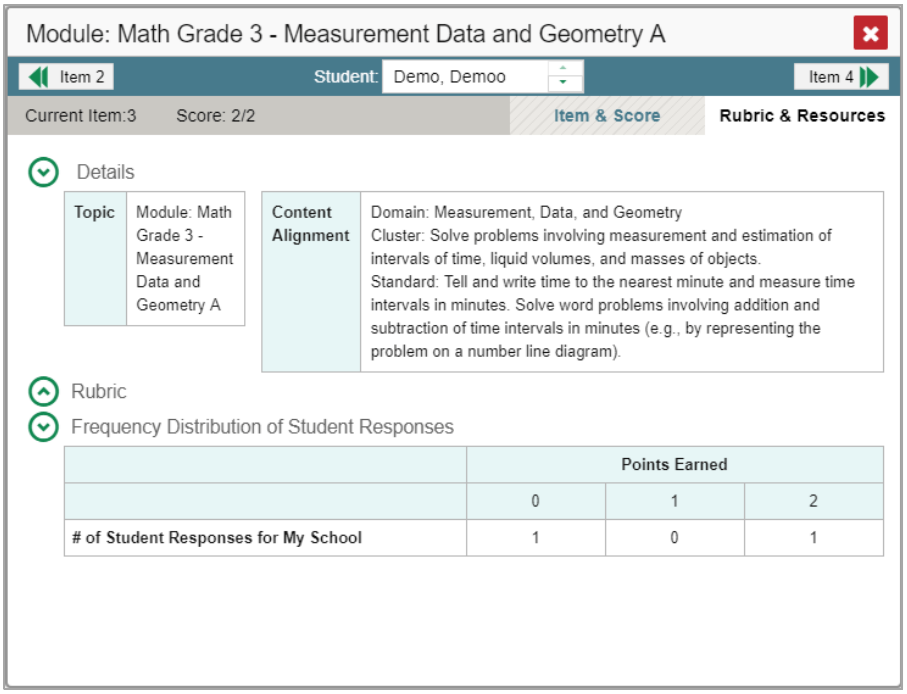
Figure 60. Item View Window: Item & Score Tab with Student Response and Scoring Assertion Table



- **Item:** Displays the item as it appeared on the assessment in the Student Testing Site. For items associated with a passage, the passage also appears.

The **Rubric & Resources** tab (see Figure 61) may include the following sections, which you can expand and collapse by clicking and , respectively.

Figure 61. Item View Window: Rubric & Resources Tab



- **Details**: May provide the following information:
  - **Topic**: Skill area to which the item belongs.
  - **Difficulty**: Indicates whether the item is intended to be easy, moderate, or difficult.
  - **Content Alignment**: Describes the standard to which the item is aligned.
- **Resources**: Provides links to any exemplars or training guides available for the item.
- **Rubric**: Displays the criteria used to score the item. This section may also include a score breakdown, a human-readable rubric, or an exemplar, which provides an example of a response for each point value.
- **Frequency Distribution of Student Responses**: The table in this section provides a breakdown of how many students in the school earned each possible point value available for a fixed-form test item.

How to View Items with and Without the Students' Visual Settings

When viewing items with students' responses, you may or may not want to see the items exactly the way the students saw them on the test. For example, some students' tests are set to use large fonts, different color contrast, or Spanish.

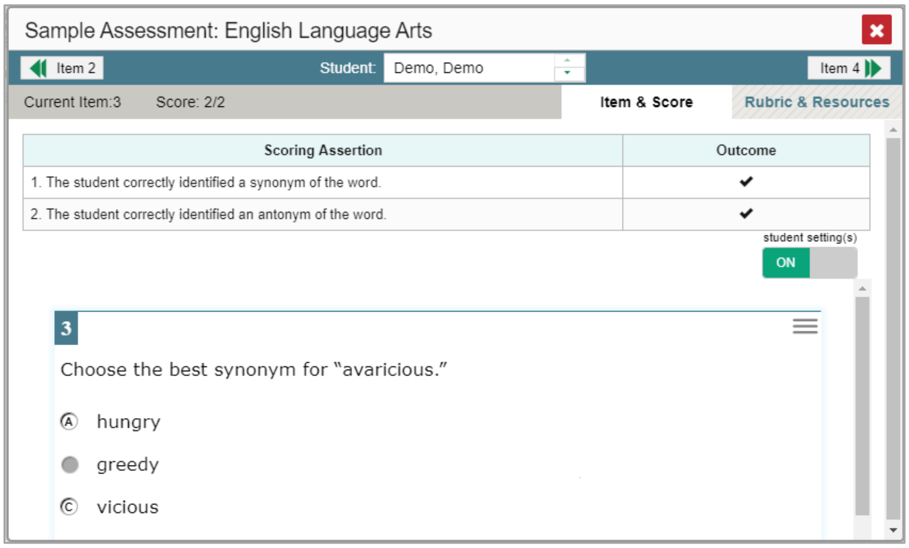
1. From the **Features & Tools** menu, select **Set Student Setting on Item View**. The ***Set Student Setting on Item View*** window appears.

2. Select **Yes** to show students' visual settings on all items or **No** to hide them.

3. Click **Save**.

You can also show or hide visual settings on a per-item basis. To do so, click the toggle at the upper right of the item you're viewing (see Figure 62). This action has no effect on your global setting.

Figure 62. Item View Window: Item & Score Tab with Student Response



What It Means When a Student Response Contains Highlighted Text

When a student's text response contains too much text copied from the item prompt and a condition code of Insufficient Original Text to Score has been applied, the copied portion is automatically highlighted.

How to Navigate to Other Items from the Item View Window

Use the buttons labeled with the previous and next item numbers at the upper corners of the ***Item View*** window.

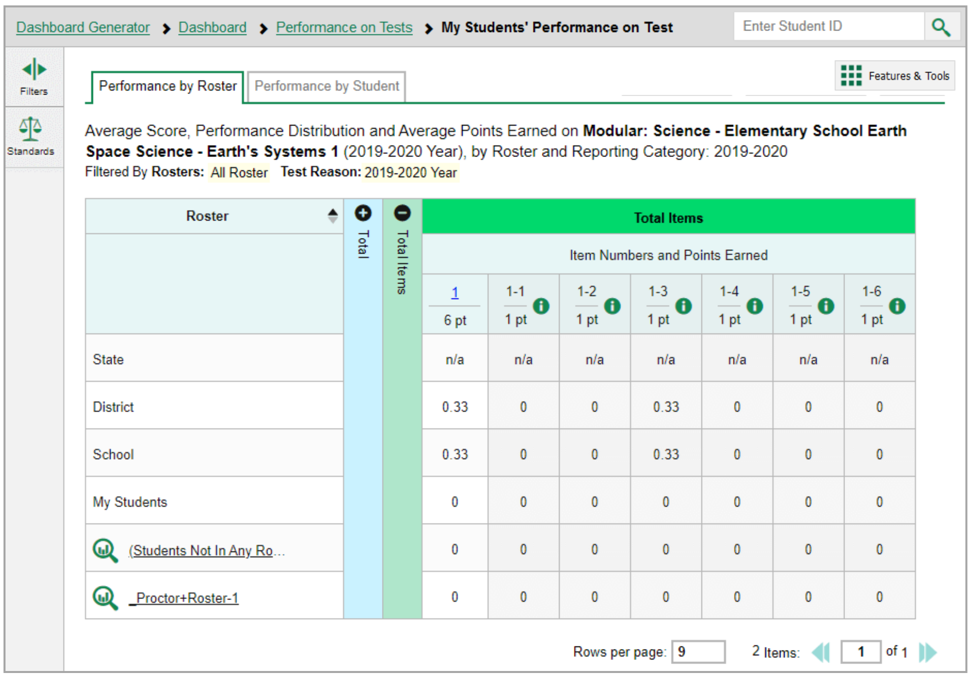
How to View Another Student's Response to the Current Item

If you have accessed the student's response from a report showing multiple students, you can click the arrows beside the *Student* field Demo Student at the top of the window. The students are listed in the same order in which they are sorted in the report.

What It Means When Items Are Labeled “1-1”, “1-2”, and So On (MSA Only)

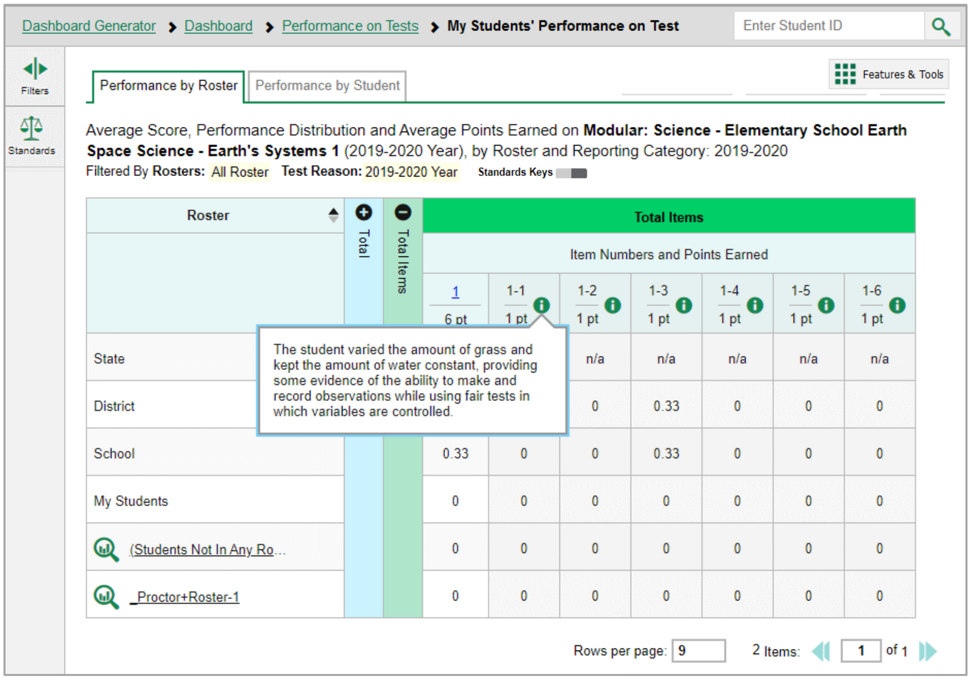
Those are sub-items belonging to an item cluster. Clusters are broken down into sub-items because they have multiple scoring assertions. Each sub-item has its own column to the right of the main item column. Sub-items are labeled “[item number]-[sub-item number]”, for example, “1-1”, “1-2”, “1-3”, as in Figure 63. This is applicable to the Montana Science tests.

Figure 63. My Students’ Performance on Test Report: Performance by Student Tab with Expanded Total Items Section



To view a scoring assertion, click the more information button to the right of the sub-item number, as in Figure 64.

Figure 64. My Students' Performance on Test Report: Performance by Student Tab with Expanded Total Items Section and Sub-Item Assertion



What It Means When an Item Score Reads "n/a"

You may sometimes see "n/a" instead of a score for an item. In some cases, the student did not respond to the item, or the item was not included in that form of the test.

How to Score Items on Interims (SBAC Only)

The Reporting System allows authorized users to score certain items on interim tests.

- Some items that require hand scoring arrive in the Reporting System without any scores. For example, all short answer items require hand scoring. If a test contains unscored items, its performance data are excluded from your reports until an authorized user scores all the unscored items in at least one opportunity of that test.

- Other items arrive in the Reporting System with automated scores suggested by the machine scoring system. Authorized users can override these scores if necessary. For example, all full write items have machine-suggested scores that can be overridden.

How to Score Unscored Items

For a student's test performance to be reported, you need to enter scores for any hand-scored item responses on that test.

When you have tests with unscored items, a **Tests To Score** notification appears in the banner.

1. In the banner, click **Tests To Score**.
2. If the ***Select School*** page appears, make a selection and click **Continue**.
3. On the scoring ***Dashboard*** (Figure 65), click the name of the test you wish to score. The ***Test Scoring*** page appears (Figure 66), displaying a list of students and items awaiting scoring for the selected test.

Figure 65. Scoring Dashboard

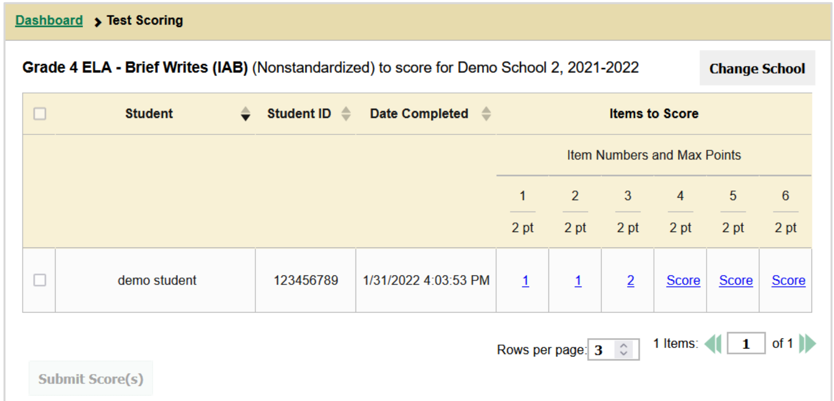
Dashboard

Change School

Assessments to score for Demo School 2, 2021-2022

Assessment Name	Test Reason	Student Count	Items to Score	Items to Submit	Last Date Completed
Grade 5 MATH - Performance Task - Turtle Habitat (IAB)	Nonstandardized	1	4	0	02/15/2022
Grade 5 ELA - Brief Writes (IAB)	Nonstandardized	1	4	2	02/15/2022
Grade 4 ELA - Performance Task (ICA)	Standardized/Benchmark	1	2	0	02/14/2022
Grade 4 ELA - Brief Writes (IAB)	Standardized/Benchmark	1	2	4	02/14/2022
Grade 3 ELA - Read Literary Texts (IAB)	Nonstandardized	1	1	0	01/31/2022

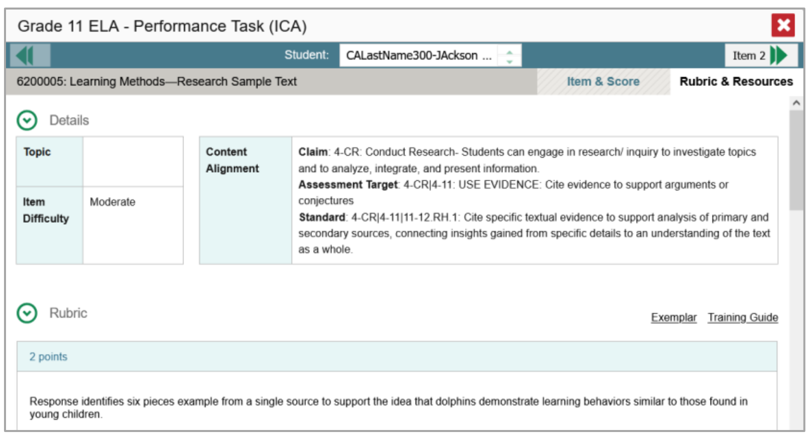
Figure 66. Test Scoring Page



4. To enter scores for an item response, click the **Score** link for the required item in the required student's row. The scoring window opens to the **Item & Score** tab.

5. *Optional*: You can open the **Rubric & Resources** tab (Figure 67) to review the item's scoring rubric and any other available resources, such as an exemplar and training guide for scoring the item.

Figure 67. Item View Window: Rubric & Resources Tab



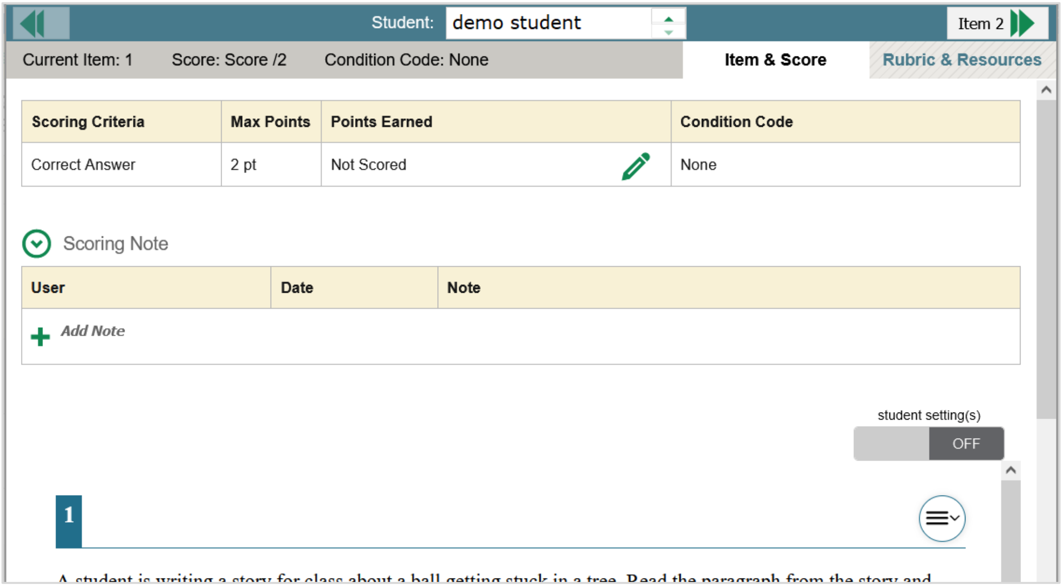
6. On the **Item & Score** tab (see Figure 68), click the edit button in the Scoring Criteria table at the top of the window.

7. Review the student's entered response and do one of the following:

   - If the student's response is scorable, select the appropriate score from the drop-down menu in the Points Earned column.

   - If the student's response cannot be scored for any reason, select the appropriate condition code from the drop-down in the Condition Code column. For more information about condition codes, see Condition Codes.

Figure 68. Item View Window: Item & Score Tab



8. If the item has multiple scoring criteria, repeat step 7 for each criterion.

   - When assigning condition codes to multiple scoring criteria, note that some related criteria may require the codes to be the same.

9. Click **Save**.

10. *Optional*: If you wish to provide an explanation for why you chose the given score or condition code, click **Add Note** in the *Scoring Note* section, then enter your comment and click **Save**.

11. To continue scoring items, do one of the following:

    - To view another unscored item for the same student, use the buttons labeled with the previous and next item numbers at the upper corners of the scoring window.
    - To view the same unscored item for another student, click the up or down arrows on the right side of the *Student* field Demo Student at the top of the window.
    - To return to the ***Scoring Mode*** window and select another item manually, close the ***Item View*** window using the close button at the upper-right corner.

12. Repeat steps 5–11 until you have entered scores for all the unscored items for the test.

13. *Optional*: If you need to change the entered score for an item response, you can select the score link for that item on the ***Test Scoring*** page and repeat steps 5–11.

Once you have entered scores for all the unscored items in a test opportunity, you may submit that opportunity for processing.

Note that for some items, two Scoring Criteria tables appear, with the top one allowing you to set scores and the bottom one displaying transformed scores (see Figure 69). When you set a score, the new scores are automatically transformed, and the transformed scores are what will appear in reports. You will need to refresh or navigate away from the item before you can view the transformed scores, and there may be a delay before they appear.

Figure 69. Item View Window: Item & Score Tab: Second Scoring Criteria Table with Transformed Scores

Transformed Points Earned and Condition Codes

Scoring Criteria	Max Points	Points Earned	Condition Code
Conventions	2	2	None
Organization, Purpose, Evidence and Elaboration	4	Condition Code Selected	Off Purpose
Overall	6	2	None

Full write items are scored using three dimensions: Conventions, Evidence/Elaboration and Organization/Purpose for a total of ten points. For test scoring purposes,

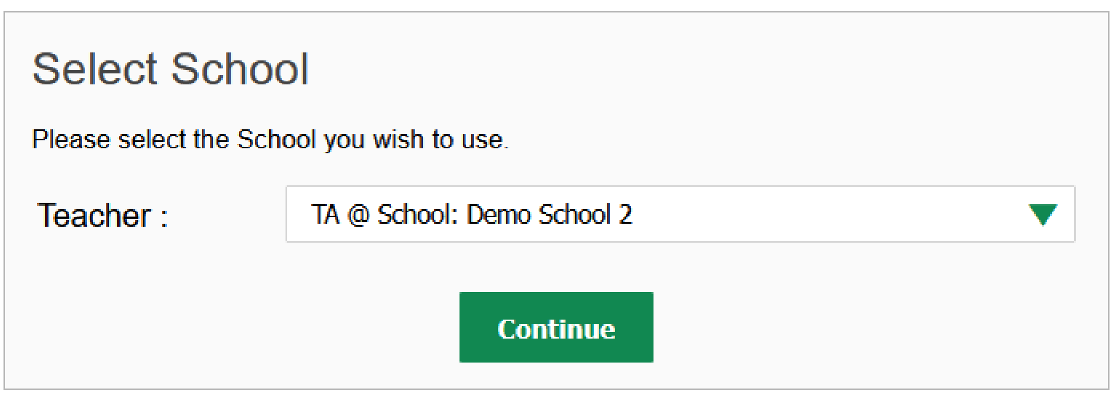
To return to reports, click the close button at the upper-right corner.

How to Change the Selected School

Scoring mode allows you to score the item responses for only one school's test opportunities at a time. If you are associated with multiple schools, you may change the selected school in order to score the test opportunities from other schools.

1. To change the selected school, click **Change School** in the top-right corner of the ***Dashboard***. The ***Select School*** page appears.

Figure 70. Select School Page



2. From the **Teacher** drop-down menu, select the school whose student responses you wish to score.

3. Click **Continue**. The ***Dashboard*** page appears, displaying the tests available for the selected school.

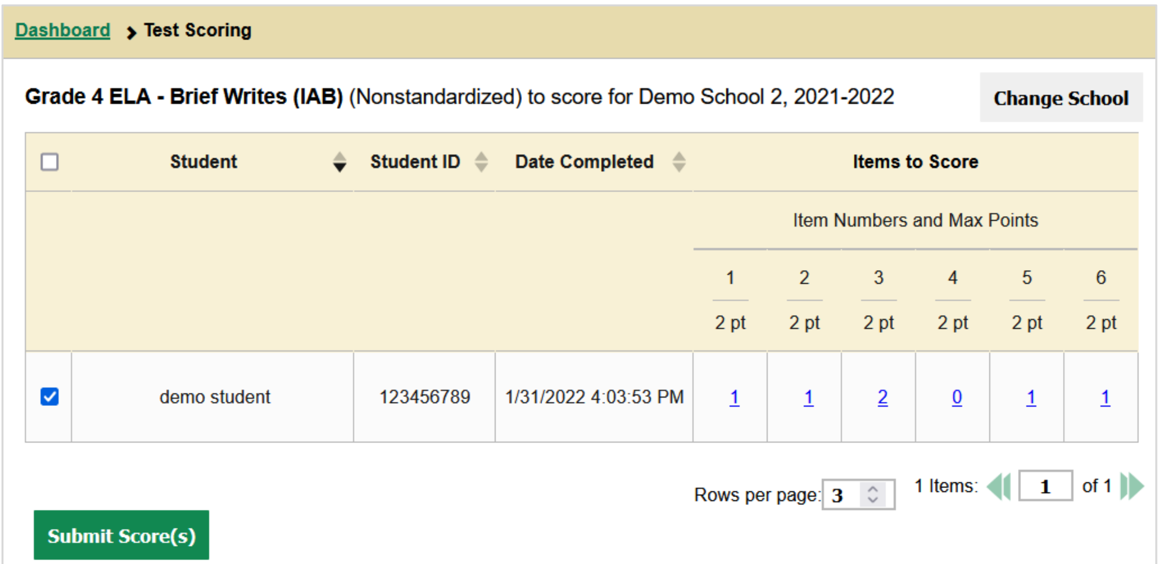
How to Submit Scored Test Opportunities for Processing

In order for a test opportunity's scores to be reported, you will need to submit that opportunity for processing. You may only submit an opportunity once you have entered scores or condition codes for every hand-scored item on the test.

1. To submit scored opportunities for processing, navigate to the ***Test Scoring*** page (see Figure 71) and mark the checkbox for each opportunity you wish to submit.

    - You can mark the checkbox in the top-left corner to select all fully scored opportunities at once.

Figure 71. Test Scoring Page: Submitting an Opportunity



2. Click **Submit Score(s)** in the bottom-left corner of the page.

3. In the confirmation window that pops up, click **Continue**. The selected opportunities will be submitted for processing and reporting and removed from scoring.

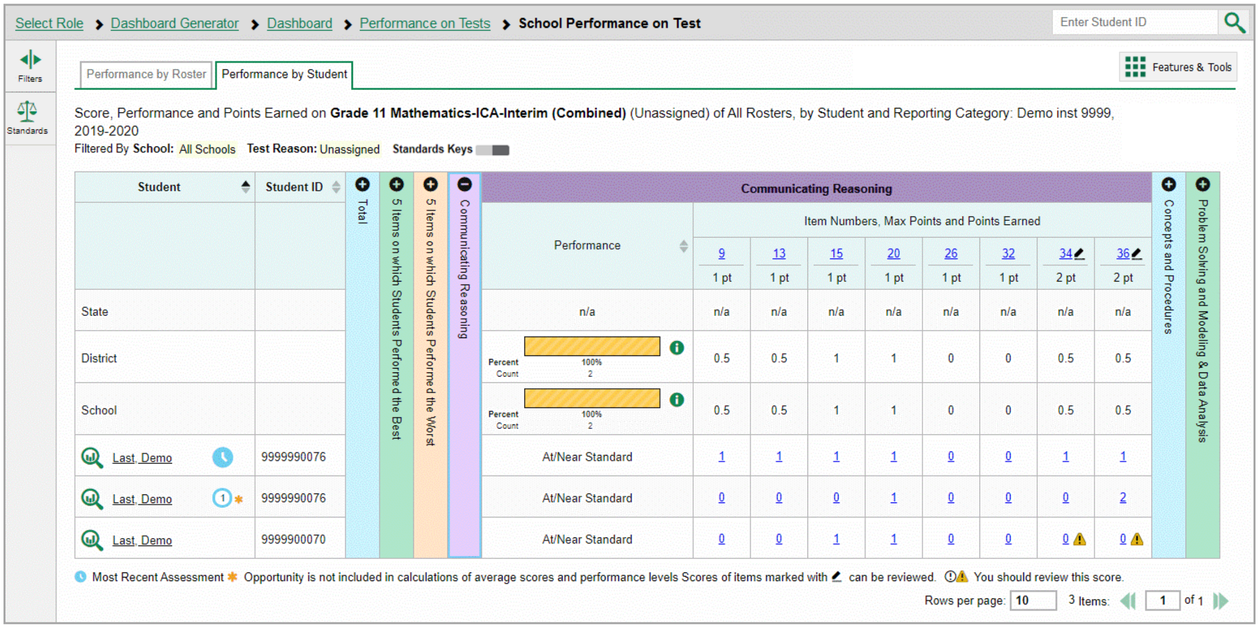
To return to reports, click the close button at the upper-right corner. You can still modify the item scores on that test directly from the reports by following the procedure in the next section (How to Modify Scores for Items).

How to Modify Scores for Items

You can modify scores for some items directly from the ***Item View*** window.

Reports display a pencil icon in the column header for each item with a modifiable score (see Figure 72). When a machine-suggested score has a low confidence level, or when a condition code of Insufficient Text or Non-Scorable Language has been assigned by machine, ⚠ displays next to the score. It is highly recommended that you review items flagged with this icon.

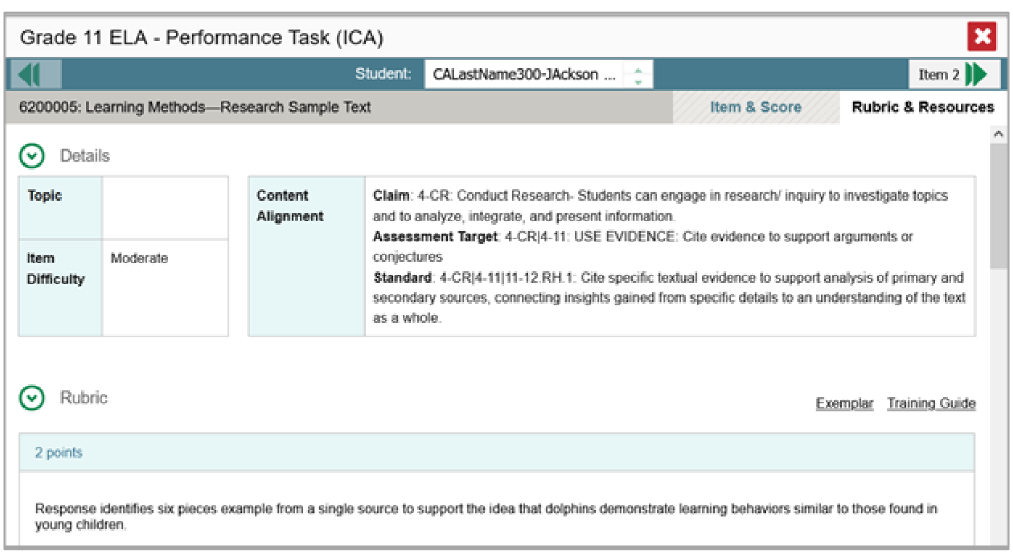
Figure 72. School Performance on Test Report: Performance by Student Tab with Expanded Reporting Category Section



1. On a report with modifiable scores, click the item score link in the student's row of the report. The ***Item View*** window opens.

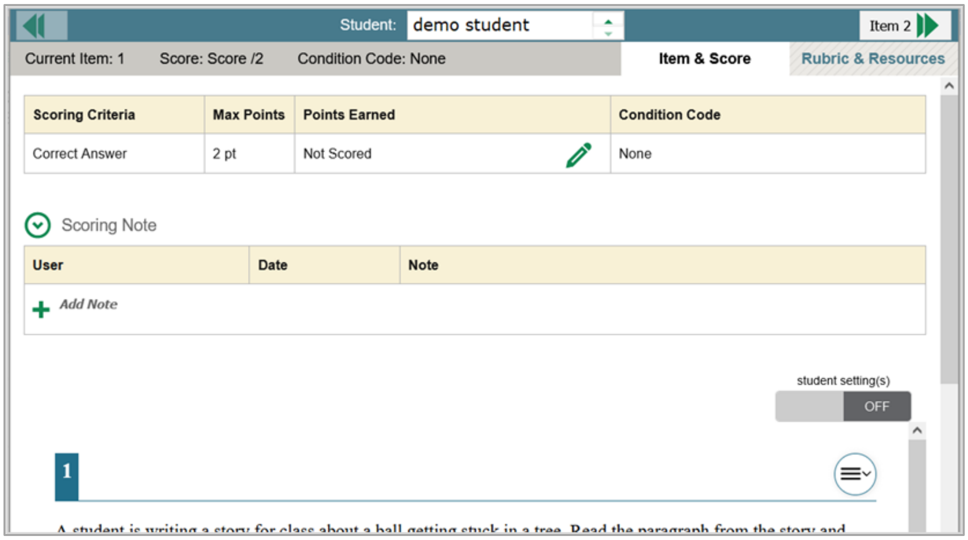
2. *Optional*: You can open the **Rubric & Resources** tab (Figure 73) to review the item's scoring rubric and any other available resources, such as an exemplar and training guide for scoring the item.

Figure 73. Item View Window: Rubric & Resources Tab



3. On the **Item & Score** tab (see Figure 74), review the student's entered response and click in the Scoring Criteria table at the top of the window. The Points Earned and Condition Code columns become editable.

Figure 74. Item View Window: Item & Score Tab



4. Do one of the following:

   - To enter a score for the response, select a numerical score from the **Points Earned** drop-down list.

   - To assign a condition code to the response, select one from the **Condition Code** drop-down list.

5. If the item has multiple scoring criteria, repeat step 4 for each criterion.

6. Click **Save**.

7. *Optional*: If you wish to provide an explanation for why you chose the given score or condition code, click **Add Note** in the *Scoring Note* section, then enter your comment and click **Save**.

8. To continue modifying scores, do one of the following:

   - To view another item for the same student, use the buttons labeled with the previous and next item numbers at the upper corners of the ***Item View*** window.

   - To view the same item for another student, use the up or down arrow buttons on the right side of the *Student* field Demo Student at the top of the ***Item View*** window.

The performance data in the test results update automatically when you close the ***Item View*** window.

Note that for some items, two Scoring Criteria tables appear, with the top one having modifiable scores and the bottom one displaying transformed scores, as in Figure 75. When you modify a score, the new scores are automatically transformed, and the transformed scores are what will appear in reports. You will need to refresh or navigate away from the item or the report before you can view the transformed scores, and there may be a delay before they appear.

Figure 75. Item View Window: Item & Score Tab: Second Scoring Criteria Table with Transformed Scores

Transformed Points Earned and Condition Codes

Scoring Criteria	Max Points	Points Earned	Condition Code
Conventions	2	2	None
Organization, Purpose, Evidence and Elaboration	4	Condition Code Selected	Off Purpose
Overall	6	2	None

Full write items are scored using three dimensions: Conventions, Evidence/Elaboration and Organization/Purpose for a total of ten points. For test scoring purposes,

How to Set Up Interim Reports to Suit Your Needs

There are three ways of setting up your interim reports that are different from summatives. You can assign test reasons to interim test opportunities, filter them by test reason, and filter them by standard.

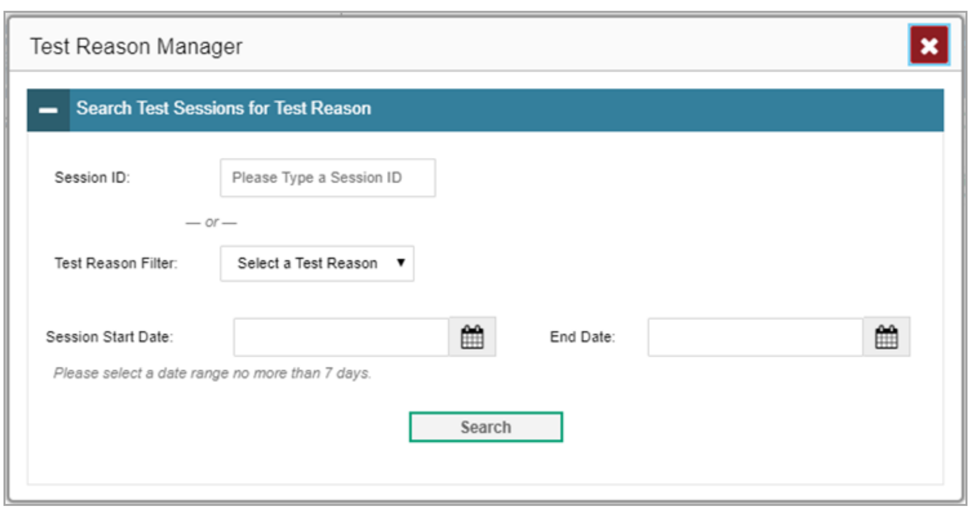
How to Assign Test Reasons (Categories) to Interim Test Opportunities

Test reasons are categories used to classify test opportunities for reporting purposes. They typically indicate the timeframe in which tests were taken, and they're a good way to organize tests into groups.

Test reasons should ideally be assigned in the Test Administration Site at the time of testing. However, you can use the Test Reason Manager in the Reporting System to assign a different test reason to an interim or benchmark test opportunity that was completed in the present school year. Summative test reasons cannot be reassigned.

1. From the **Features & Tools menu**, select **Manage Test Reasons**. The ***Test Reason Manager*** window opens (see Figure 76).

Figure 76. Test Reason Manager Window

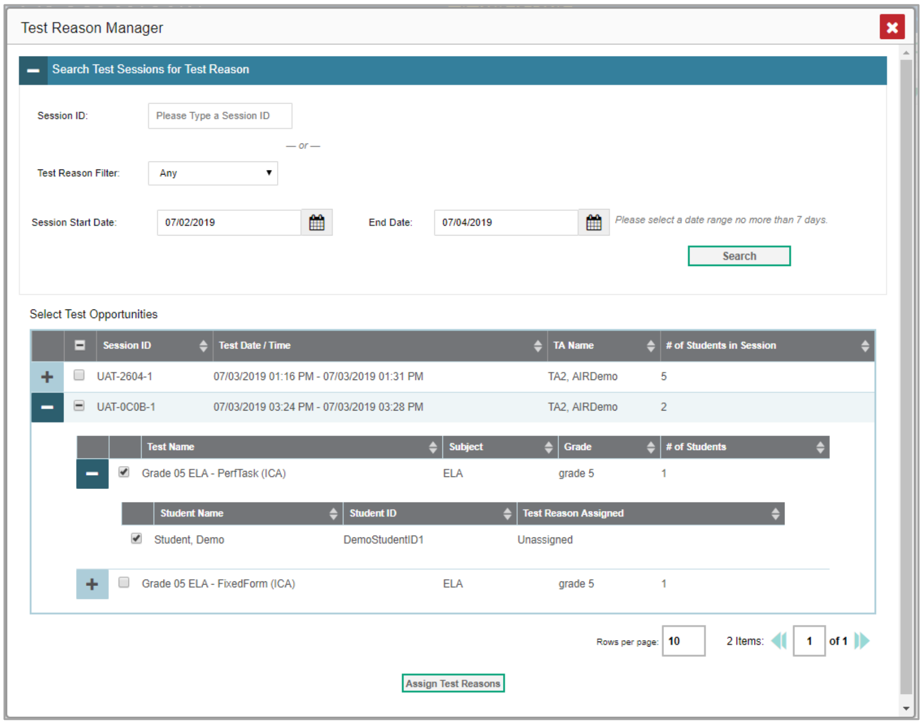


2. To search for the test opportunities you wish to categorize, do either of the following (see Figure 76):

- In the *Session ID* field, enter the session ID in which the opportunities were completed in TDS.
- Select the test reason associated with the opportunities you want to edit. Then select a range of dates during which the test session was administered. The date range cannot exceed seven days.

3. Click **Search**.
4. A list of retrieved test sessions appears in the section *Select Test Opportunities* (see Figure 77). You can click the + buttons to expand the list of tests in each session and the list of students who took each test (that is, individual test opportunities). To navigate through a long list, use the controls in the upper-right and lower-right corners.

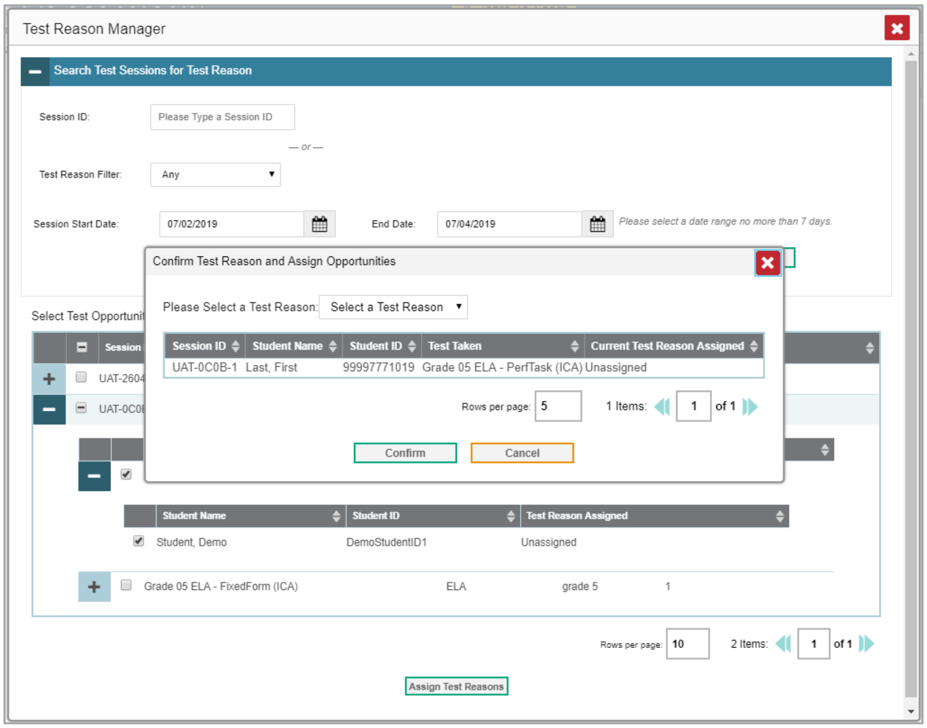
Figure 77. Test Reason Manager Window: Select Test Opportunities



5. Mark the checkboxes for each session, test, or opportunity that you wish to assign to a test reason.
6. Click **Assign Test Reasons** below the list of retrieved sessions.

7. In the window that appears (see Figure 78), select a new test reason to assign to the selected opportunities and click **Confirm**.

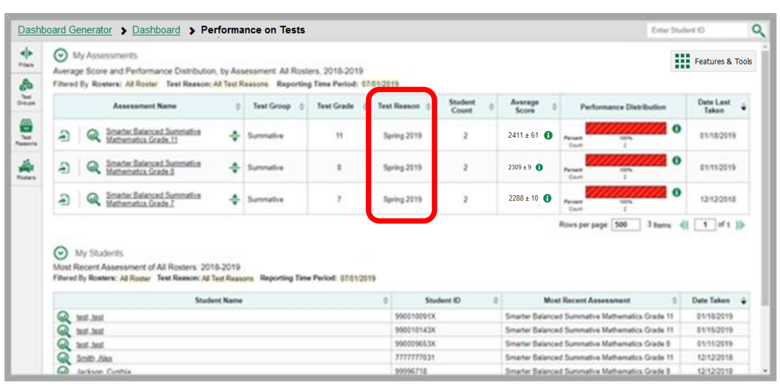
Figure 78. Confirm Test Reason and Assign Opportunities Window



How to Filter by Test Reason (Category)

Test reasons (shown in Figure 79) are categories used to classify test opportunities for reporting purposes. They typically indicate the timeframe in which interim and benchmark tests were taken, and they can be a good way to focus on specific groups of tests. For summative assessments, test reasons are simply test windows and are not useful.

Figure 79. Teacher View: Performance on Tests Report



When your test opportunities have test reasons, you can filter reports by a single test reason. For example, you may want to filter by Fall and look at ELA performance, then filter by Spring and see if students have improved on ELA material. If you don't filter, you'll see data for all different test reasons.

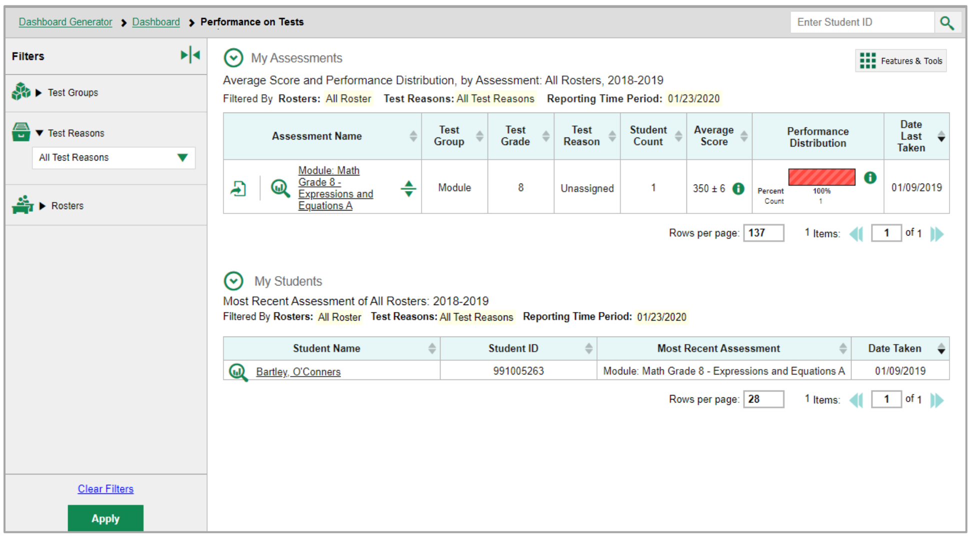
This will allow you to compare multiple test reasons side by side rather than a single test reason. You may find reports easier to understand when you're viewing only a single test reason.

The **Test Reasons** filter is available on the dashboards and Performance on Tests reports for teachers as well as for school- and district-level users.

1. On the left side of the dashboard or Performance on Tests report, click either the **Filters** panel expand button or the **Test Reasons** button. The **Filters** panel expands (see Figure 80).

2. Make a selection from the drop-down list in the **Test Reasons** section.

Figure 80. Teacher View: Performance on Tests Report with Expanded Filters Panel



3. Click **Apply**. The report updates to show only data for that test reason.

4. *Optional*: To revert all filters to their defaults, open the **Filters** panel again and click **Clear Filters**. Click **Apply**. Filters will also revert when you log out, switch user roles, or switch systems.

All the reports accessible from this page will be filtered the same way.

The row of filter details below the table header shows the test reason selected, if any.

How to Filter Item-Level Data on Interims by Standards and Clusters of Standards

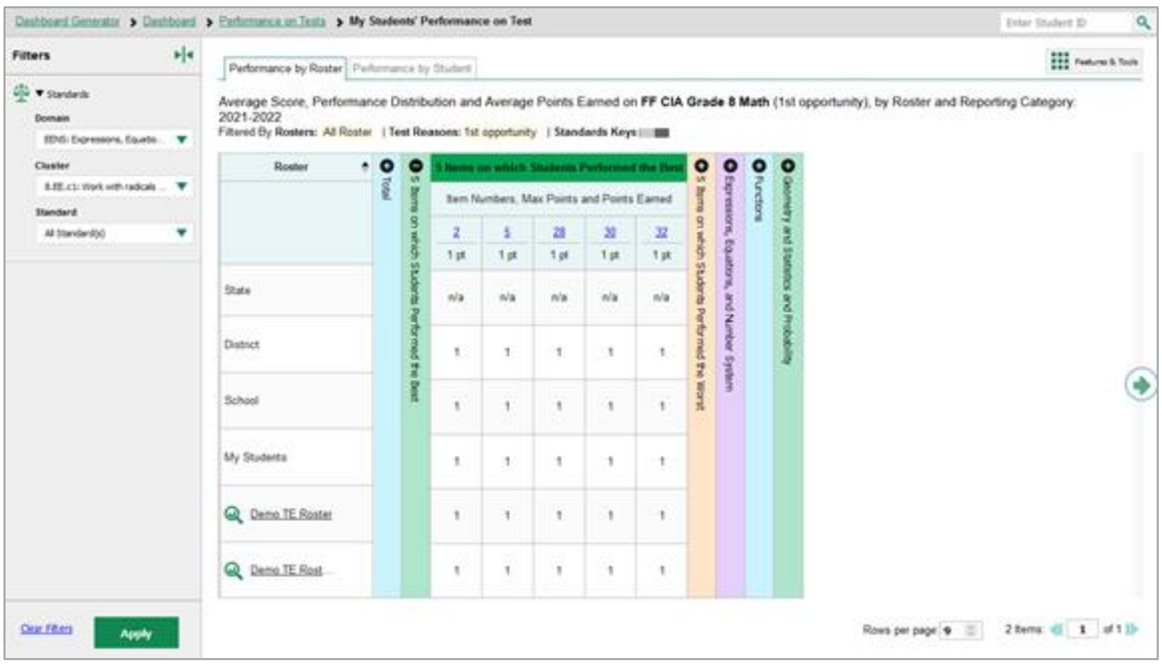
An educational standard, sometimes called an assessment target, describes the skill the item measures. An example of a math standard is "At later grades, determine conditions under which an argument does and does not apply. (For example, area increases with perimeter for squares, but not for all plane figures.)"

You may want to see how your students performed on a particular standard or cluster of standards. In certain reports, you can filter by the standard to which items are aligned. That way you can view your students' performance in just one area of skill. Then you can switch filters to compare it with their performance in another skill. If you don't filter by standard, the reports will show results for all standards by default. You may find that switching between different sets of standard data and comparing them helps you understand students' abilities better.

Standard filters are available in any report showing item-level data. The available standards vary by assessment.

1. On the left side of the page, click either the **Filters** panel expand button or the **Standards** button . The **Filters** panel expands (see Figure 81).
2. Use the drop-down list in the **Standards** section to select a cluster. An additional drop-down list appears.
3. *Optional*: Keep making selections from the drop-down lists as they appear.

Figure 81. My Students' Performance on Test Report: Performance by Roster Tab with Expanded Filters Panel



4. Click **Apply**. The affected report updates to show only the items that belong to the selected cluster or standard.
5. *Optional*: To revert all filters to their defaults, open the filters panel again and click **Clear Filters**. Click **Apply**. Filters will also revert when you log out, switch user roles, or switch systems.

All the reports accessible from this page will be filtered the same way.

The row of filter details below the table header specifies the standards selected, if any.

How to Export and Print Data on Interims

You can export or print any report you see in the Reporting System. Some reports on individual tests can be exported directly from the Performance on Tests report. You may want to export or print to save a snapshot of data to consult later, or to share data. Different options will be available depending on the report you are viewing. Some interim and benchmark reports can be exported with item-level data.

How to Export or Print a Report You're Viewing

1. Select the print button from the **Features & Tools** menu, or in a Longitudinal Report window, from the upper-right corner. If there are multiple report tables on the page, multiple print buttons will appear.

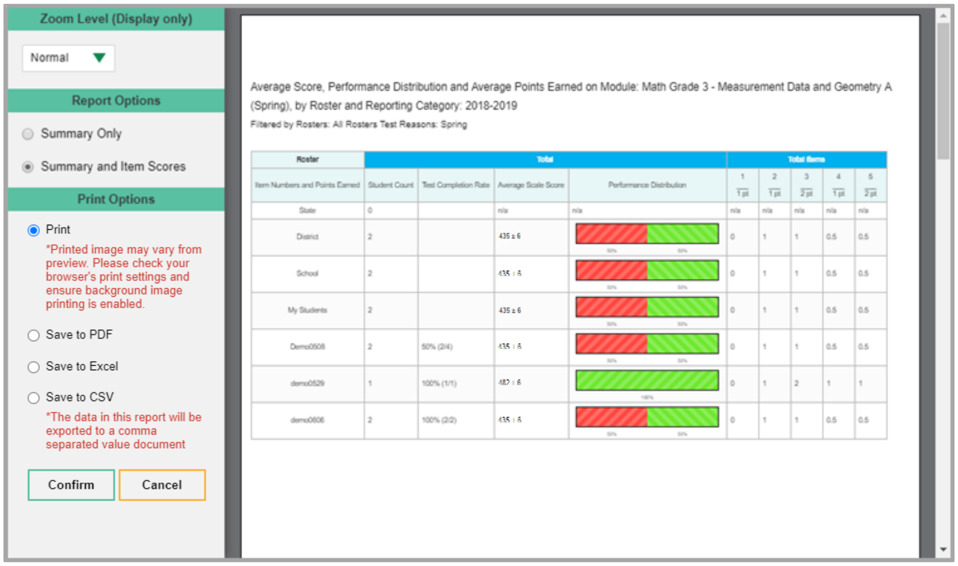
   A print preview page opens (see Figure 82).

   - To zoom in on the print preview, use the drop-down list under the *Zoom Level (Display only)* section. This setting affects the preview only.

Figure 82. Print Preview Page

2. If an aggregate report provides data for individual items, the *Report Options* section appears. Select either **Summary Only** or **Summary and Item Scores**. If you select the latter option, as in Figure 83, the printed report includes data for the individual assessment items. Printouts of the Student Performance on Test report always include item data if available.

Figure 83. Print Preview Page with Summary and Item Scores Option Selected



3. To print the report, select the **Print** radio button. To download it, select **Save to PDF**, **Save to Excel** (.xlsx), or **Save to CSV** (comma-separated values).

- *Optional*: If a printout or PDF is for a particular student, you can mark the **Include Items and Responses (takes extra time)** checkbox. The resulting report includes the actual items and the student's responses.
- If you selected **Save to PDF**, choose an option from the **Page Layout** drop-down list that appears.

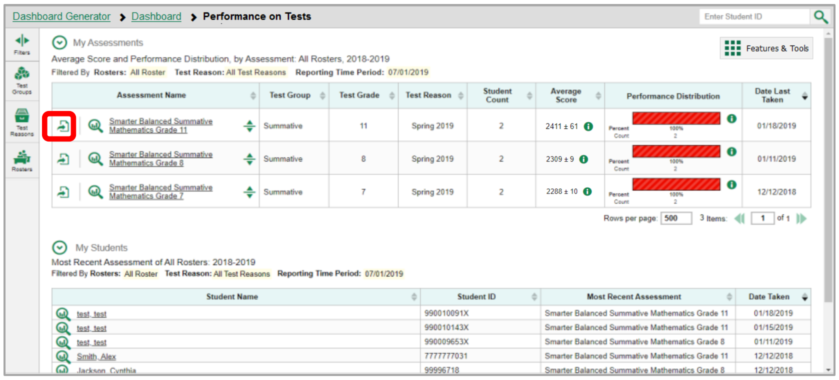
4. Click **Confirm**.

If you saved the report as a PDF, Excel, or CSV, the ***Secure Inbox*** window appears, displaying the generated report.

How to Export an Assessment Report Directly from the Performance on Tests Report

1. Click  to the left of the name of the assessment whose report you wish to export (see Figure 84).

Figure 84. Teacher View: Performance on Tests Report

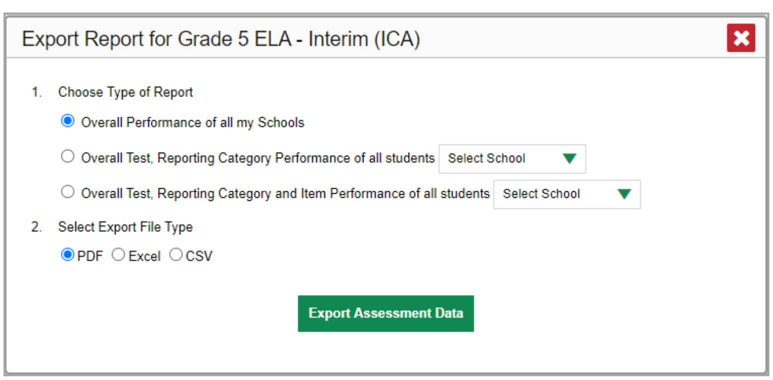


The ***Export Report*** window opens (see Figure 85 and Figure 86). The options in this window vary according to your user role.

2. Select which report to export for the assessment.

- **District-level users:**
  - To export the district test results, mark the **Overall Performance of all my Schools** radio button.
  - To export school test results (excluding data for individual items), mark the **Overall Test & Reporting Category Performance of all my Students for [School Name]** radio button, then select a school from the drop-down list.
  - To export school test results (including data for individual items), mark the **Overall Test, Reporting Category and Item Performance of all my Students for [School Name]** radio button, then select a school from the drop-down list.

Figure 85. District-Level User View: Export Report Window

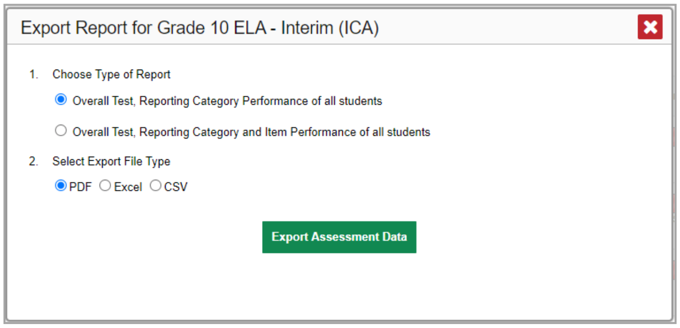


- **School-level users and teachers:**

- To export results for all your associated students (excluding data for individual items), mark the **Overall Test, Reporting Category Performance of all students** radio button.

- To export results for all your associated students (including data for individual items), mark the **Overall Test, Reporting Category and Item Performance of all students** radio button.

Figure 86. Teacher View: Export Report Window



3. Choose from the **PDF**, **Excel**, and **CSV** formats.
4. Click **Export Assessment Data**. A confirmation window appears.
5. Click **Yes** to export or **No** to return to the ***Export Report*** window. When you've exported a file, the ***Secure Inbox*** window appears with the generated file available for download.

Appendix

Appendix sections are alphabetized for your convenience.

C

Class (Roster) Management

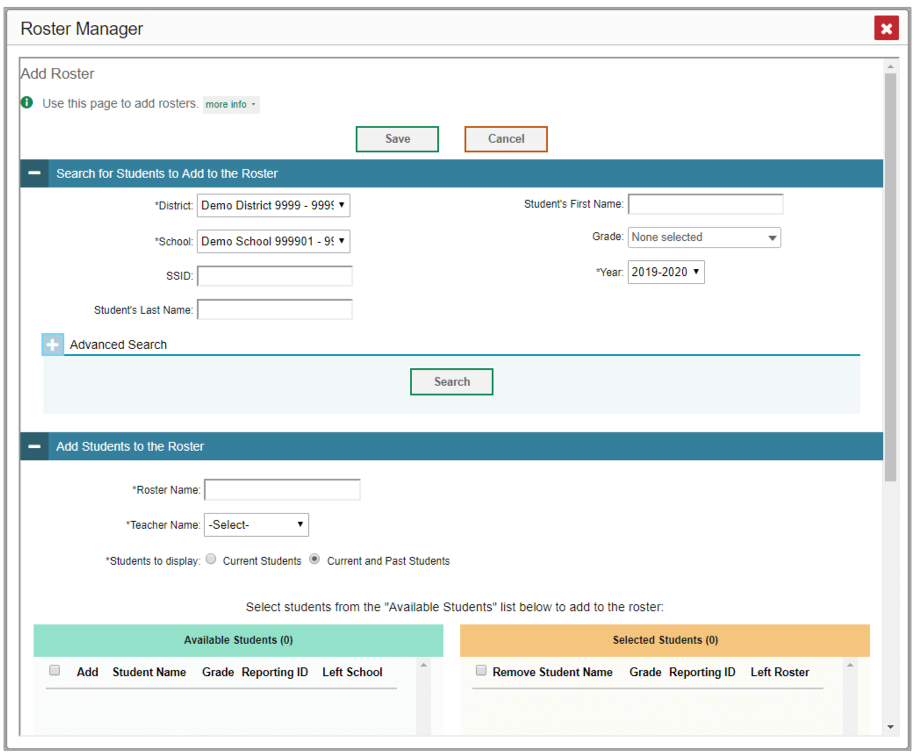
Teachers, school-level users, and district-level users can add, edit, and delete classes (rosters). Classes are a great way to organize students, allow teachers to view their students' performance, and allow other users to compare the performance of different classes.

How to Add a Class (Roster)

You can create new classes (rosters) from students associated with your school or district.

1. From the **Features & Tools menu**, select **Add Roster**. The ***Roster Manager*** window appears, showing the Add Roster form (see Figure 87).

Figure 87. Roster Manager Window: Add Roster Form



2. In the *Search for Students to Add to the Roster* panel (see Figure 87), do the following:

6. If you are a district-level user, then in the **School** drop-down list, select the school for the roster.

   b. *Optional*: In the *SSID*, *Student's First Name*, and/or *Student's Last Name* fields, enter information about a particular student you want to add.

7. *Optional*: In the **Enrolled Grade** drop-down list, select the grade levels for the students in the roster.

8. *Optional*: In the *Advanced Search* panel (see Figure 88), select additional criteria:

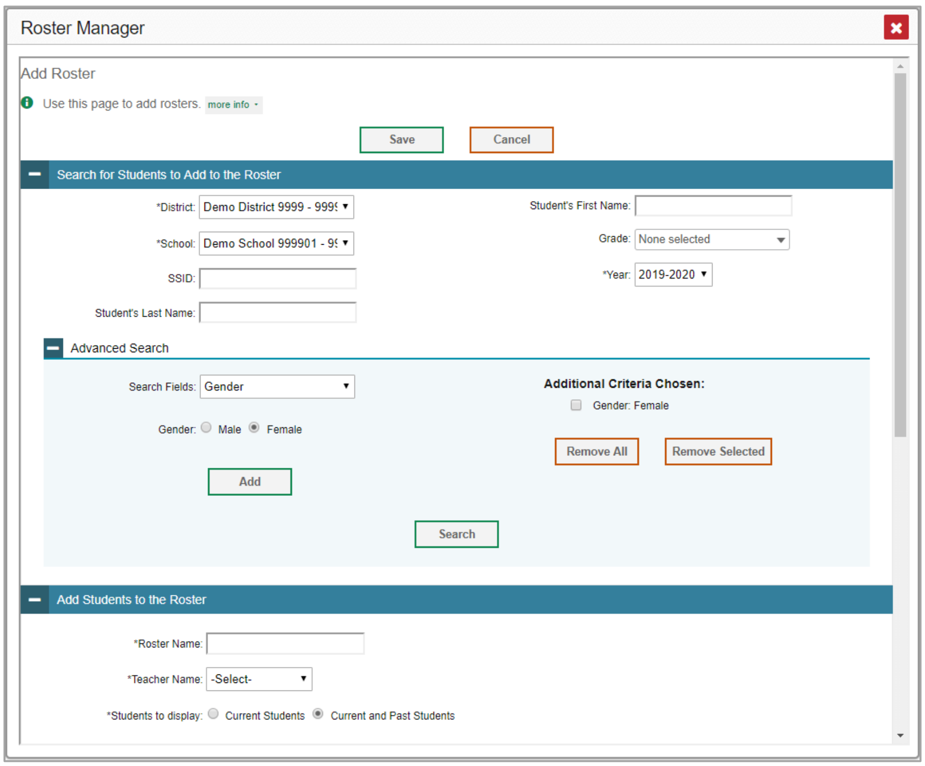
    i. From the **Search Fields** drop-down list, select a criterion type. A set of related criteria for that criterion type appear.

    ii. In the related fields, select the additional criteria.

    iii. Click **Add**.

    iv. *Optional*: To remove the added criteria, mark the checkboxes for those criteria and click **Remove Selected**. To remove all additional criteria, click **Remove All**.

Figure 88. Roster Manager Window: Add Roster Form with Advanced Search Panel in Use



9. Click **Search**. The *Add Students to the Roster* panel shows settings for the roster, a list of retrieved students (*Available Students*), and a blank *Selected Students* list.

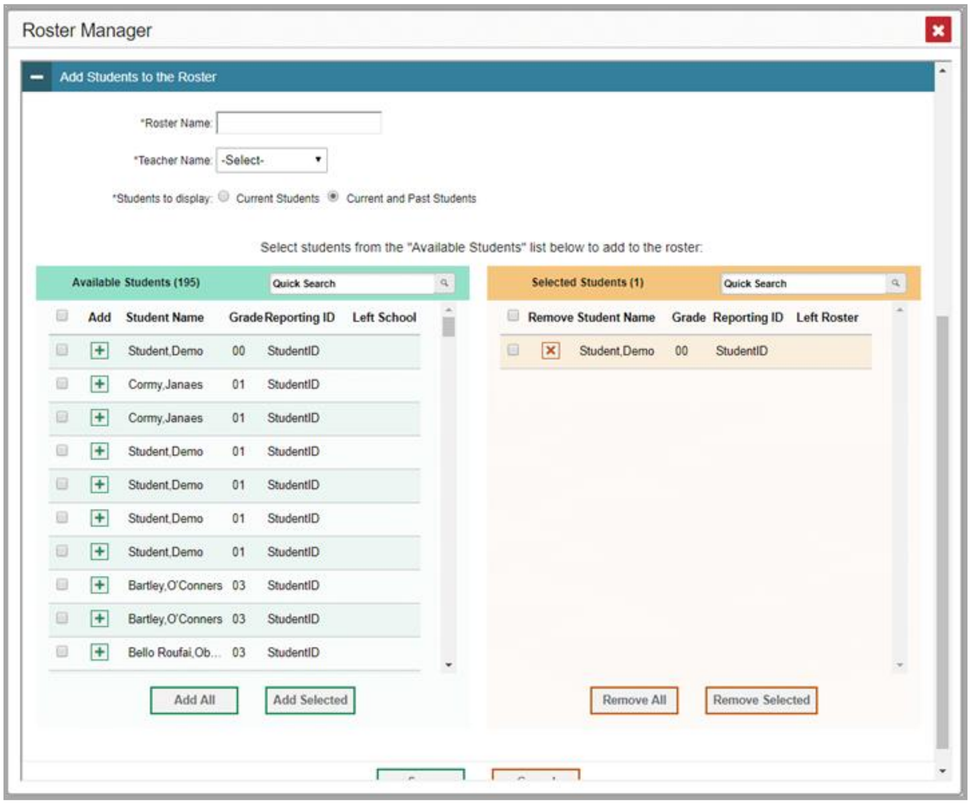
3. In the *Add Students to the Roster* panel (see Figure 89), do the following:

    a. In the *Roster Name* field, enter the roster name.

10. From the **Teacher Name** drop-down list, select a teacher.

11. *Optional*: To include former students in the Add Roster form, mark the **Current and Past Students** radio button. The *Available Students* list will include students who have left the selected school.

Figure 89. Roster Manager: Add Roster Form Scrolled Down to Add Students to the Roster Panel



12. To add students, do one of the following in the list of available students:

- To move one student to the roster, click [+] beside that student's name.
- To move all the students in the *Available Students* list to the roster, click **Add All**.
- To move selected students to the roster, mark the checkboxes for the students you want to add, then click **Add Selected**.

13. To remove students, do one of the following in the list of students in this roster:

- To remove one student from the roster, click [X] beside that student's name.
- To remove all the students from the roster, click **Remove All**.
- To remove selected students from the roster, mark the checkboxes for the students you want to remove, then click **Remove Selected**.

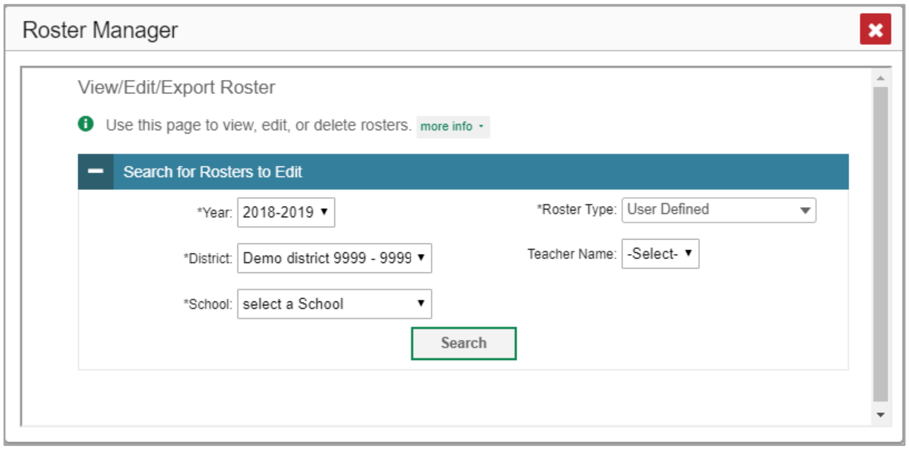
4. Click **Save**, and in the affirmation dialog box click **Continue**.

How to Modify a Class (Roster)

You can modify a class (roster) by changing its name, changing its associated teacher, adding students, or removing students.

1. From the **Features & Tools menu**, select **View/Edit Roster**. The ***Roster Manager*** window appears, showing the View/Edit/Export Roster form (see Figure 90).

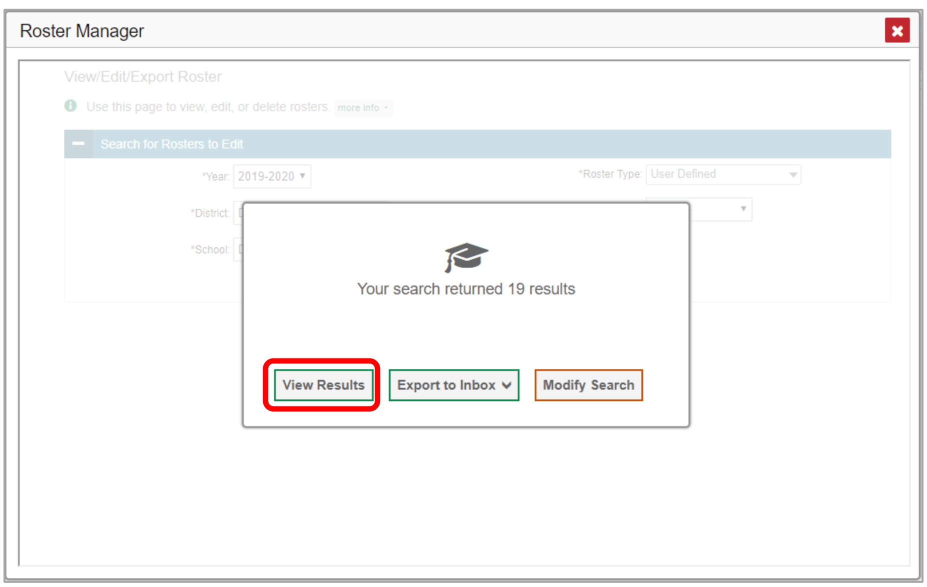
Figure 90. Roster Manager Window: View/Edit/Export Roster Form



2. In the *Search for Rosters to Edit* panel (see Figure 90), select the school year, school, and roster type for the roster you wish to edit. Optionally, select a teacher.

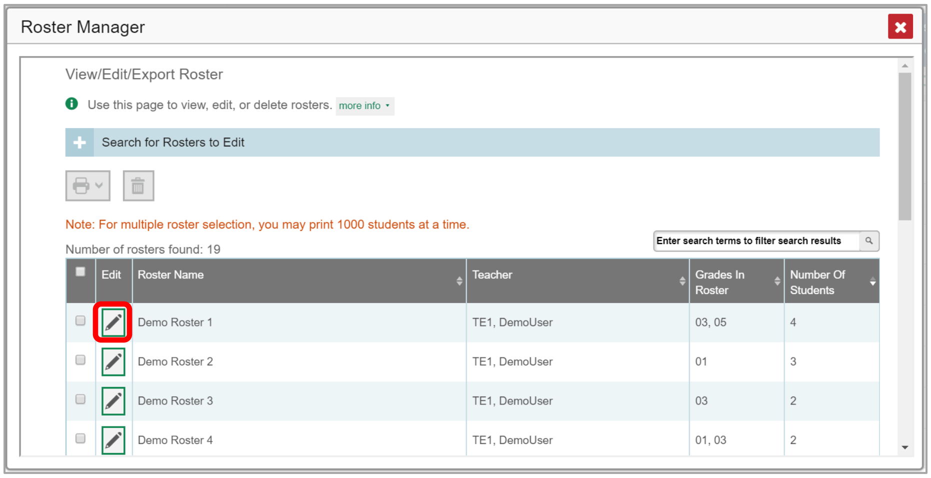
3. Click **Search**. A search results pop-up appears (see Figure 91). Click **View Results** to view the results in your browser.

Figure 91. Roster Manager Window: Search Results Pop-Up



4. A list of retrieved rosters is generated (see Figure 92).

Figure 92. Roster Manager Window: View/Edit/Export Roster Form Showing Retrieved Rosters



5. In the list of retrieved rosters, click [edit icon] for the roster whose details you want to view. The ***View/Edit/Export Roster*** window opens.

6. *Optional*: To find students to add to the roster, use the *Search for Students to Add to the Roster* panel as follows:

a. If you are a district-level user, then in the **School** drop-down list, select the school for the roster.

b. *Optional*: In the SSID, *Student's First Name*, and/or *Student's Last Name* fields, enter information about a particular student you want to add.

14. *Optional*: In the Enrolled **Grade** drop-down list, select the grade levels for the students in the roster.

15. *Optional*: In the *Advanced Search* panel, select additional criteria:

i. From the **Search Fields** drop-down list, select a criterion type. A set of related criteria for that criterion type appear.

ii. In the related fields, select the additional criteria.

iii. Click **Add**.

iv. *Optional*: To remove the added criteria, mark the checkboxes for those criteria and click **Remove Selected**. To remove all additional criteria, click **Remove All**.

16. Click **Search**. The *Add Students to the Roster* panel shows settings for the roster, a list of retrieved students (Available Students), and a blank Selected Students list.

7. Scroll down to view the *Add Students to the Roster* panel, as in Figure 93.

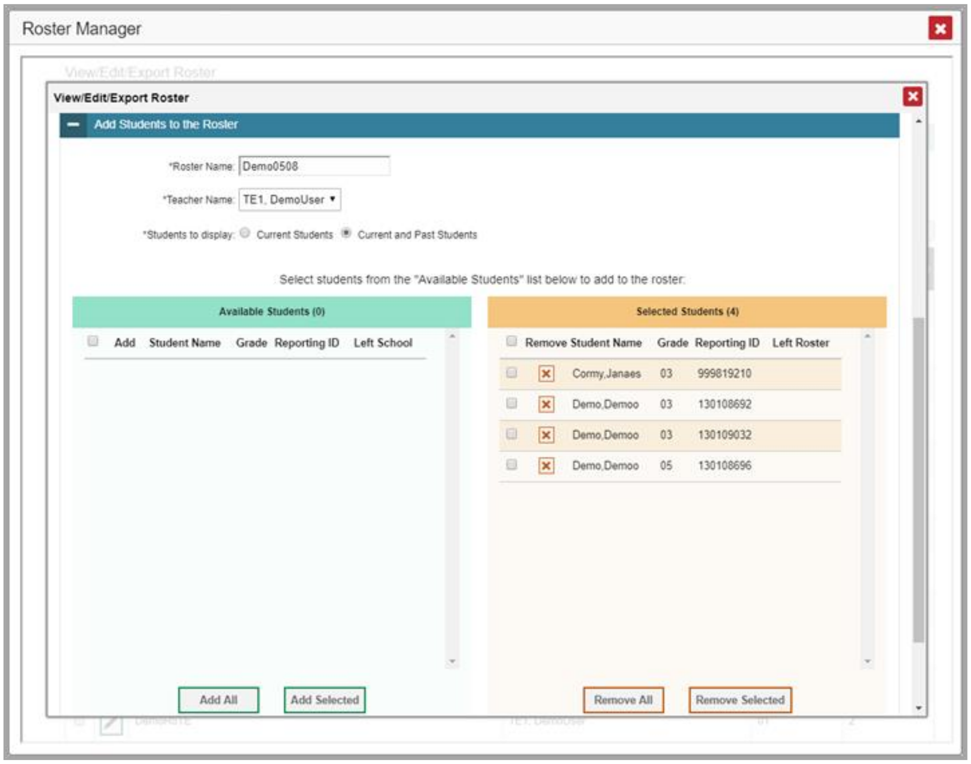


Figure 93. Roster Manager Window: View/Edit/Export Roster Form Scrolled Down to the Add Students to the Roster Panel

8. *Optional*: In the *Add Students to the Roster* panel, do the following:

a. In the *Roster Name* field, enter a new name for the roster.

17. From the **Teacher Name** drop-down list, select the roster's new teacher.

18. *Optional*: To include former students in the Edit Roster form, mark the **Current and Past Students** radio button. The *Available Students* list will include students who have left the selected school, while the *Selected Students* list will include students who have left the roster.

19. To add students, do one of the following in the list of available students:

- To move one student to the roster, click beside that student's name.
- To move all the students in the *Available Students* list to the roster, click **Add All**.
- To move selected students to the roster, mark the checkboxes for the students you want to add, then click **Add Selected**.

20. To remove students, do one of the following in the list of students in this roster:

- To remove one student from the roster, click beside that student's name.

- To remove all the students from the roster, click **Remove All**.
- To remove selected students from the roster, mark the checkboxes for the students you want to remove, then click **Remove Selected**.

9. At the bottom of the page, click **Save**, and in the affirmation dialog box click **Continue**.

How to Upload Classes (Rosters)

If you have many classes (rosters) to create, it may be easier to perform those transactions through file uploads. This task requires familiarity with composing comma-separated value (CSV) files or working with Microsoft Excel.

1. From the **Features & Tools menu**, select **Upload Rosters**. The ***Roster Manager*** window appears, showing the Upload Rosters: Upload page (see Figure 94).

Figure 94. Roster Manager Window: Upload Rosters: Upload Page

2. On the Upload Rosters: Upload page (see Figure 94), click **Download Templates** in the upper-right corner and select the appropriate file type (either **Excel** or **CSV**).

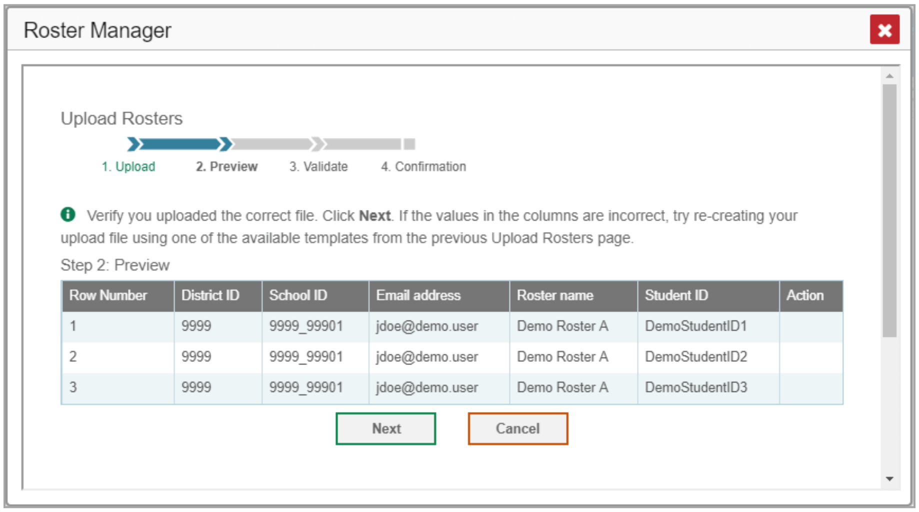
3. Open the template file in a spreadsheet application.

4. Fill out the template and save it.

5. On the Upload Rosters: Upload page, click **Browse** and select the file you created in the previous step.

6. Click **Next**. The Upload Rosters: Preview page appears (see Figure 95). Use the file preview on this page to verify you uploaded the correct file.

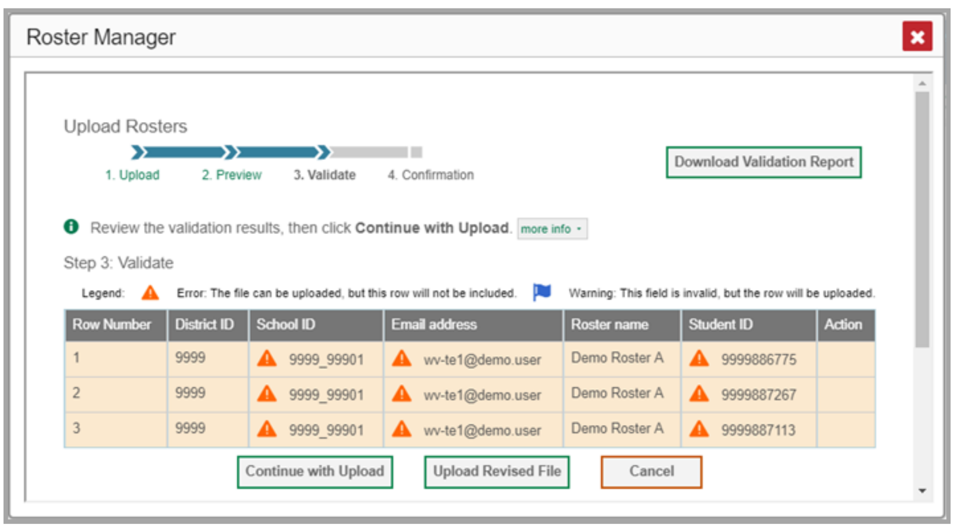
Figure 95. Roster Manager Window: Upload Rosters: Preview Page



7. Click **Next** to validate the file.

Any errors or warnings are displayed on the Upload Rosters: Validate page (see Figure 96). If a record contains an error, that record will not be included in the upload. If a record contains a warning, that record will be uploaded, but the field with the warning will be invalid.

Figure 96. Roster Manager Window: Upload Rosters: Validate Page



- *Optional*: Click the error and warning icons in the validation results to view the reason a field is invalid.
- *Optional*: Click **Download Validation Report** in the upper-right corner to view a text file listing the validation results for the upload file.

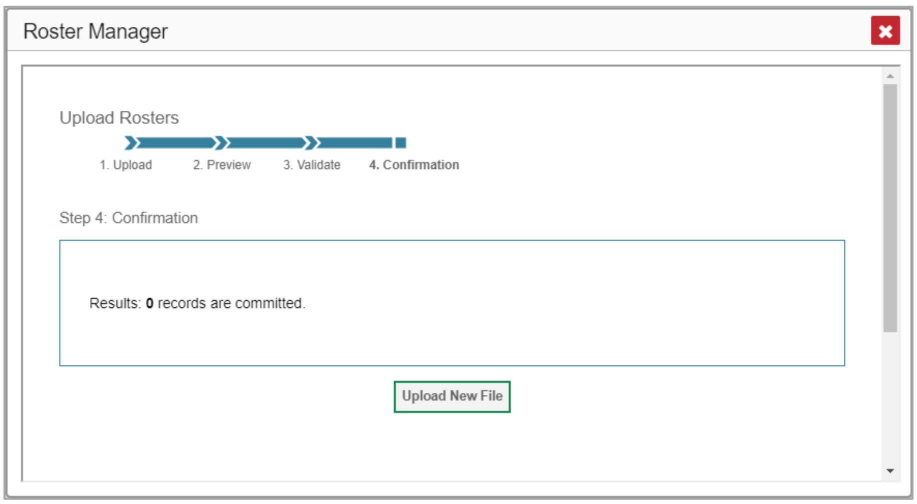
If your file contains a large number of records, the Reporting System processes it offline and sends you a confirmation email when it's complete. While the Reporting System is validating the file, do not press **Cancel**, as some records may have already started processing.

8. Do one of the following:

- Click **Continue with Upload** at the bottom of the page. The Reporting System commits those records that do not have errors. If there are too many errors, you won't be able to do this.
- Click **Upload Revised File** at the bottom of the page to upload a different file. Follow the prompts on the Upload Revised File page to submit, validate, and commit the file.

The Confirmation page appears (see Figure 97), displaying a message about how many records (rows) were committed.

Figure 97. Upload Rosters: Confirmation Page



9. *Optional*: To upload another roster file, click **Upload New File**.

Table 2 provides the guidelines for filling out the Roster template that you can download from the Upload Roster page.

Table 2. Columns in the Rosters Upload File

Element	Description	Valid Values
District ID*	District associated with the roster.	District ID that exists in TIDE. Up to 20 characters.
School ID*	School associated with the roster.	School ID that exists in TIDE. Up to 20 characters. Must be associated with the district ID. Can be blank when adding district-level rosters.
Email Address*	Email address of the teacher associated with the roster.	Email address of a teacher existing in TIDE.
Roster Name*	Name of the roster.	Up to 20 characters.
SSID*	Student's unique identifier within the district.	Up to 30 alphanumeric characters.

Element	Description	Valid Values
ACTION	Action to be taken on the student, either adding them to or deleting them from the roster. If blank, the student will be added.	`Add` or `Delete`.

*Required field.

Condition Codes

Table 3 provides an overview of the various condition codes that may be entered for a machine- or hand-scored item when a traditional score cannot be entered for the student's response.

Table 3. Condition Codes

Source of Code	Condition Code	Description
Human	Blank	• The student did not enter a response.
Human	Insufficient Text	• The student has not provided a meaningful response. Some examples: ▪ Random keystrokes ▪ Undecipherable text ▪ "I hate this test" ▪ "I don't know", "IDK" ▪ "I don't care" ▪ "I like pizza!" (in response to a reading passage about helicopters) ▪ Response consisting entirely of profanity • For ELA ICA Full Writes, use the "Insufficient Text" code for responses described above and also if ▪ The student's original work is insufficient to make a determination whether the student is able to organize, cite evidence/elaborate, and use conventions as defined in the rubrics. ▪ The response is too brief to make a determination regarding whether it is on purpose or on topic.
Human	Non-Scorable Language	• ELA/literacy: Language other than English. • Mathematics: Language other than English or Spanish.

Source of Code	Condition Code	Description
Human	Off Purpose	• For ELA ICA Full Writes only: ▪ A writing sample will be judged off purpose when the student has clearly not written to the purpose designated in the task. ▪ An off-purpose response addresses the topic of the task but not the purpose of the task. ▪ Note that students may use narrative techniques in an explanatory essay or use argumentative/persuasive techniques to explain, for example, and still be on purpose. ▪ Off-purpose responses are generally developed responses (essays, poems, etc.) clearly not written to the designated purpose.
Human	Off Topic	• For ELA ICA Full Writes only: ▪ A writing sample will be judged off topic when the response is unrelated to the task or the sources or shows no evidence that the student has read the task or the sources (especially for informational/explanatory and opinion/argumentative). ▪ Off-topic responses are generally substantial responses.
Machine	Blank	• The student did not enter a response.
Machine	Insufficient Text (Duplicated Text)	• The response contains a significant amount of text repeated over and over.
Machine	Insufficient Text (Too Few Words)	• The response contains too few words to be considered a valid attempt.
Machine	Insufficient Text (Copied Text from the Prompt)	• The response is largely composed of text copied from the prompt.
Machine	Insufficient Text (Refused to Answer)	• The response is a refusal to respond, in a form such as “idk” or “I don’t know.”
Machine	Non-Specific	• This condition code is assigned to machine-scored responses when TDS identifies that the response requires a condition code but cannot determine which specific condition code it requires.
Machine	Non-Scorable Language (Spanish Response)	• The response is in Spanish.

Source of Code	Condition Code	Description
Machine	Non-Scorable Language (Uninterpretable Language)	• The response is in a language other than English or Spanish.

H

Help

The Reporting System includes an online user guide.

How to Access the Online User Guide

In the banner, click **Help**. The guide opens in a pop-up window, showing the help page specific to the page you're on. For example, if you click **Help** while on the dashboard, you'll see the Overview of the Dashboard page.

I

Inbox

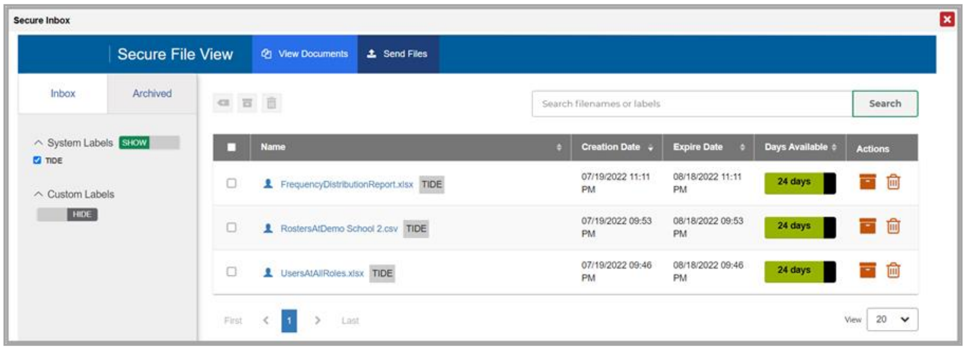
The Reporting System allows you to access a Secure Inbox feature that is integrated with other online assessment systems, such as TIDE, and accessible from your portal. The Inbox serves as a central repository for secure documents uploaded by administrators (such as state personnel) or shared between users, files exported by users, and hotline alerts.

Each user's Secure Inbox is personal to them and not shared among other users. Users can easily manage the files in their Inbox. The files are categorized into different tabs to allow users to view non-archived and archived files. Users can also search for files by keyword. Files are listed in the order in which they were created. The file creation and file expiration dates appear, if applicable, and the number of days remaining until a file expires is also displayed. By default, files are available for 30 days after being created. Users can archive or delete files as needed. Users can also share files by sending them to other users' Inboxes.

How to Access and Manage Files in the Inbox

1. In the banner, click **Inbox**. The ***Secure Inbox*** window appears (see Figure 98). By default, the Inbox window displays the **View Documents** tab.

Figure 98. Secure Inbox Window: View Documents Tab: Inbox Sub-Tab



2. Choose either of the available tabs (see Figure 98):

   - **Inbox**: Displays all files except those that have been archived. Includes columns for Creation Date, Expire Date, and Days Available.

   - **Archived**: Displays files that have been archived. Includes the same columns as the main **Inbox** tab.

3. *Optional*: To filter the files displayed, enter a search term in the text box in the upper-right corner and click . The search applies to both filenames and labels.

4. *Optional*: To hide or display system labels, click the System Labels toggle.

5. *Optional*: To hide files with a particular system label, clear the checkbox for that label.

6. *Optional*: To hide or display custom labels, click the Custom Labels toggle.

7. *Optional*: To hide files with a particular custom label, clear the checkbox for that label.

8. *Optional*: Do one of the following:

   - To download a file, click the name of the file.

   - To apply a custom label, follow these instructions:

     – To create a new custom label, mark the checkbox for any file, click the label button , enter a new custom label in the text box, and click **Save New Label**. Then apply it as described below.

     – To apply a custom label to a file, mark the checkbox for that file, click the label button , mark the checkbox for that label, and click **Apply Label**.

   - To archive a file, click .

   - To unarchive a file, click . The file is moved back to the main Inbox.

- To delete a file, click .

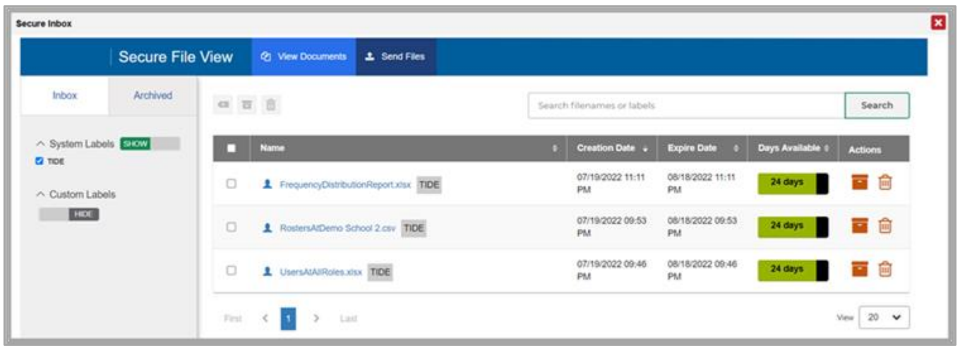
You cannot delete or archive secure documents uploaded to the Inbox by admin users.

How to Use the Inbox to Send Files to Other Users' Inboxes

You can send a file or files to individual recipients by email address or to groups of recipients by user role.

1. From the banner, select **Inbox**. The ***Secure Inbox*** page appears (see Figure 99). By default, the ***View Documents*** tab displays.

Figure 99. Secure Inbox Window: View Documents Tab: Inbox Sub-Tab



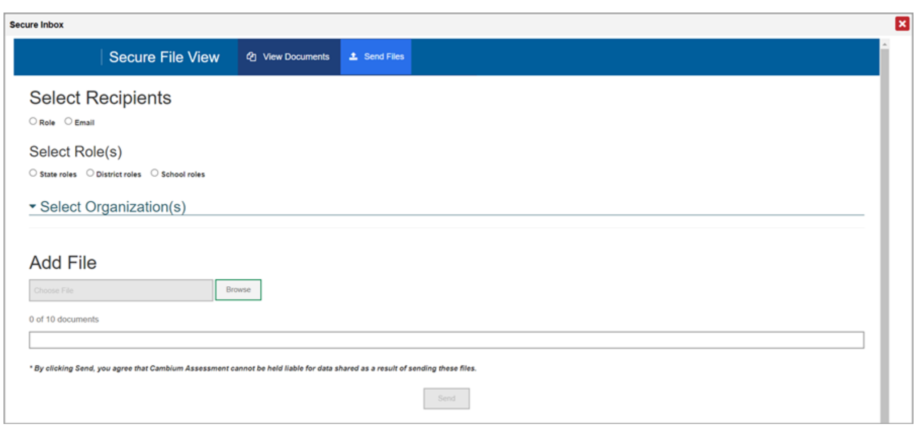
2. Select the **Send Files** tab. The ***Send Files*** page appears (see Figure 100).

3. In the *Select Recipients* field, do one of the following:

   - Select **Role** to send a file or files to a group of users by user role.
   - Select **Email** to send a file or files to a single recipient by email address.

   If you select **Email**, skip to step 7.

Figure 100. Secure Inbox Window: Send Files Tab



4. In the *Select Role(s)* field, select the role group to which you want to send a file or files. A drop-down list appears.

5. From the drop-down list, select the role(s) to which you want to send a file or files. You can choose **Select all** to send a file or files to all roles in the selected role group.

6. From the *Select Organization(s)* drop-down lists, select organizations that will receive the file(s) you send. These drop-down lists adhere to the user role hierarchy. For example, district-level users will be able to filter at their role level and below.

7. If you selected **Role** in step 3, skip this step. If you selected **Email**, enter the email address of the recipient to whom you wish to send a file or files.

8. To select a file or files to send, in the *Add File* field, select **Browse**. A file browser appears.

9. Select the file(s) you wish to send. You may send up to 10 files at once.

10. Select **Send**.

L

Login Process

This section describes how to log in to the Reporting System.

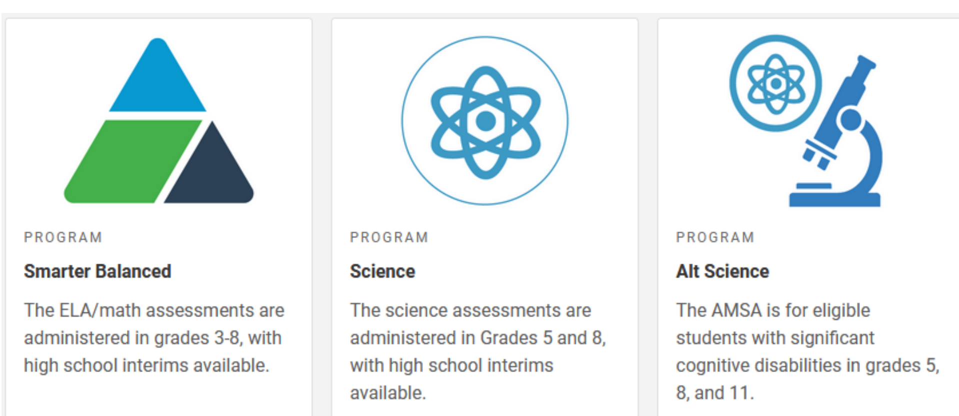
Do not share your login information with anyone. All Assessment Program systems provide access to student information, which must be protected in accordance with federal privacy laws.

How to Log In to the Reporting System

1. Navigate to the portal.

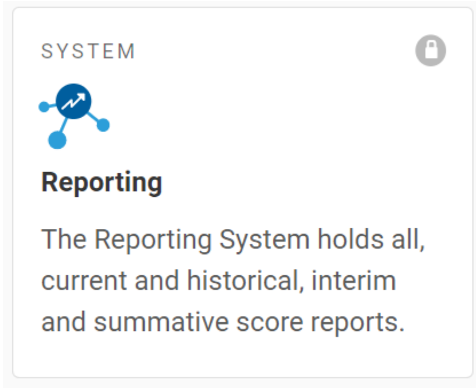
2. Select your assessment program from the cards displayed (see Figure 101).

Figure 101. Assessment Program Cards



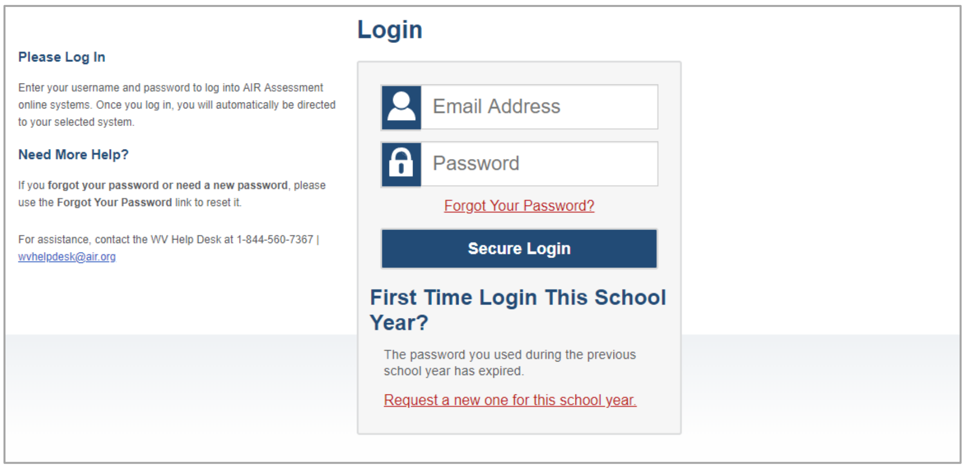
3. Click the **Reporting** card (see Figure 102). The login page appears.

Figure 102. Reporting Card



4. On the login page (see Figure 103), enter the email address and password you use to access all CAI systems.

Figure 103. Login Page

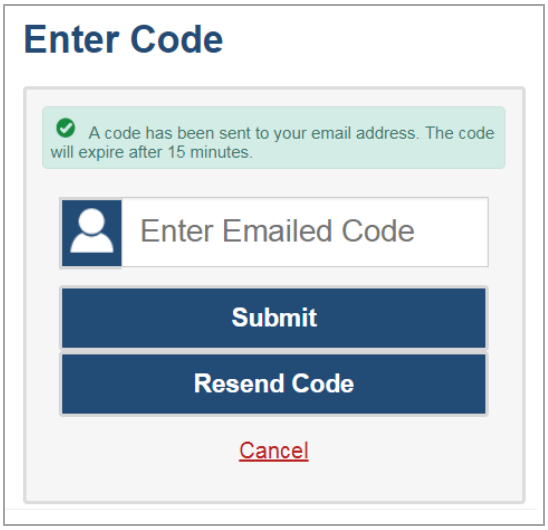


5. Click **Secure Login**.

   b. If the Enter Code page appears (see Figure 104), an authentication code is automatically sent to your email address. You must enter this code in the *Enter Emailed Code* field and click Submit within 15 minutes.

– If the authentication code has expired, click **Resend Code** to request a new code.

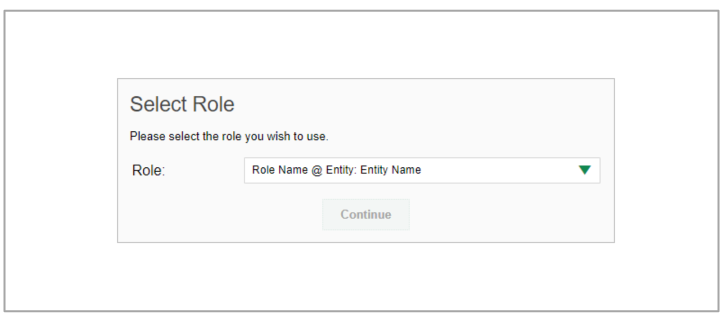
Figure 104. Enter Code Page



6. If the ***Terms and Conditions*** page appears, you should review the terms on this page and click **Accept** to proceed.

7. If your account is associated with multiple institutions, you are prompted to select a role, as in Figure 105. From the **Role** drop-down list, select the role and institution combination you wish to use. You can also change your institution after logging in.

Figure 105. Select Role Window



The Dashboard Generator for your user role appears.

How to Set or Reset Your Password

Your username is the email address associated with your account in TIDE. When you are added to TIDE, you receive an activation email containing a temporary link to the ***Reset Your Password*** page (see Figure 106). To activate your account, you must set your password within 15 minutes.

All users are required to do a one-time password reset at the beginning of every school year, for security purposes.

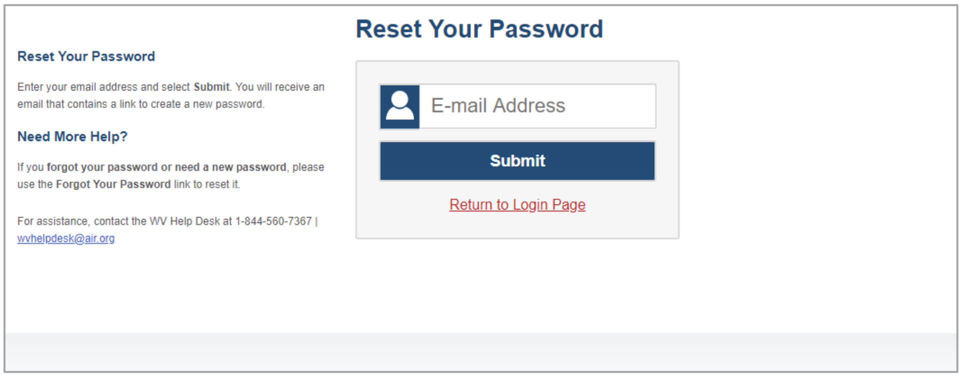
- If your first temporary link expired:

  In the activation email you received, click the second link provided and request a new temporary link.

- If you forgot your password:

  On the ***Login*** page, click **Forgot Your Password?** and enter your email address in the *E-mail Address* field. Click **Submit**. You will receive an email with a new temporary link to reset your password.

Figure 106. Reset Your Password Page



- If you did not receive an email containing a temporary link or authentication code:

  Check your spam folder to make sure your email program did not categorize it as junk mail. If you still do not see an email, contact your System Test Coordinator to make sure you are listed in TIDE.

- Additional help:

  If you are unable to log in, contact your Helpdesk for assistance. You must provide your name and email address.

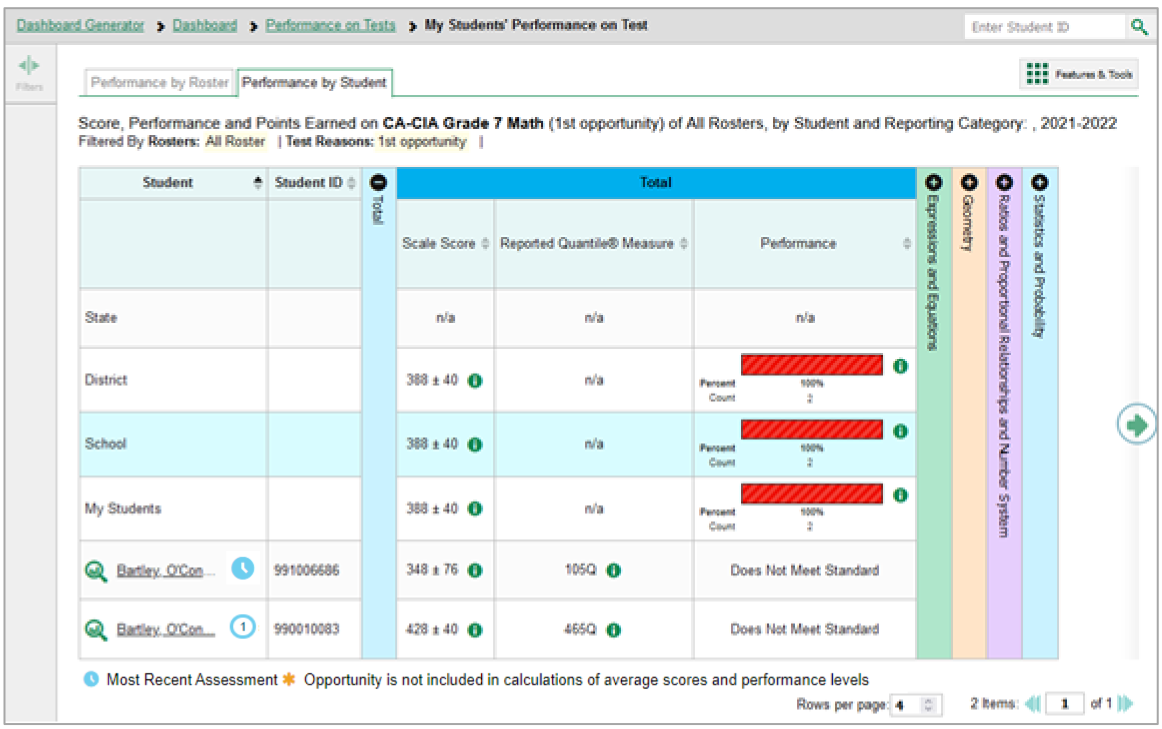
M

Multiple Interim Test Opportunities

Sometimes interim test results will include multiple rows for the same student.

When a student completes multiple test opportunities for a single assessment, as in Figure 107, reports display a row of data for each opportunity. A clock icon appears next to the most recent opportunity. Previous opportunities are marked with numbers (1), starting from the earliest test taken. An asterisk icon ✱ indicates that an opportunity is not included in calculations of average scores or performance distributions.

Figure 107. My Students' Performance on Test Report: Performance by Student Tab



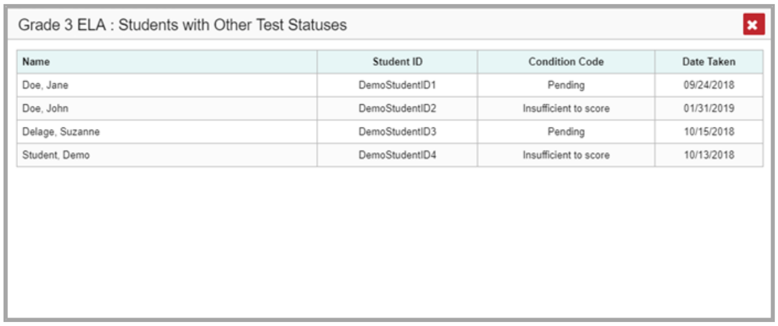
N

Non-Scorable Test Opportunities

The reports in Reporting do not include data for non-scorable test opportunities. A student's test opportunity cannot be scored when it has a test status of "Expired" or "Invalidated", or when it includes blank or empty reporting categories (reporting categories without items). If a test opportunity is non-scorable, a notification appears below the report for that assessment.

You can click **More Info** on the notification to view the ***Students with Other Test Statuses*** window (see Figure 108). This window lists the students who have non-scorable test opportunities for the given assessment, as well as the status code and completion date for each.

Figure 108. Students with Other Test Statuses Window


Grade 3 ELA : Students with Other Test Statuses

Name	Student ID	Condition Code	Date Taken
Doe, Jane	DemoStudentID1	Pending	09/24/2018
Doe, John	DemoStudentID2	Insufficient to score	01/31/2019
Delage, Suzanne	DemoStudentID3	Pending	10/15/2018
Student, Demo	DemoStudentID4	Insufficient to score	10/13/2018

P

Performance Data

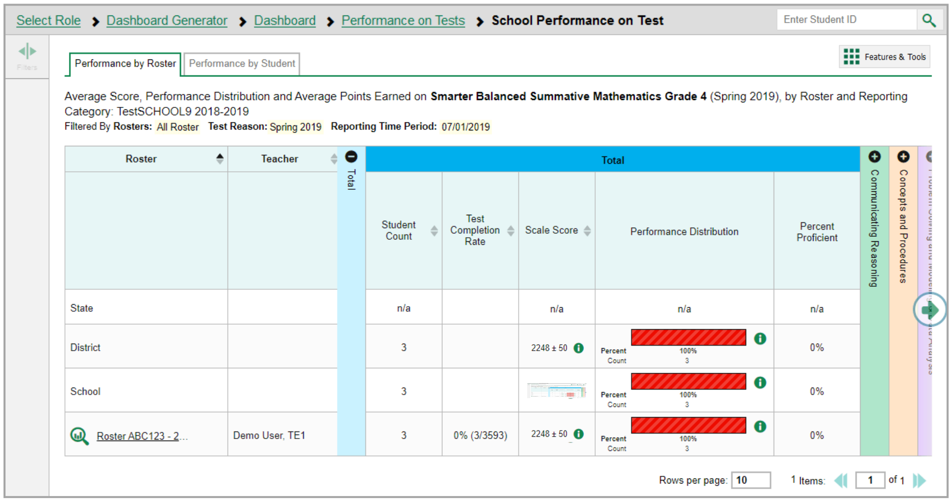
Depending on the test, a report may display different kinds of performance data:

- Score data:
  - Scale scores.
  - Raw scores, which may be in the form of percentages or fractions.
- Standard error: After an individual student's score, you may see a number with "±" before it. This is the standard error of measurement (SEM). A student's score is best interpreted when recognizing that the student's knowledge and skills fall within a score range and not just a precise number. As an example, if a student receives a test score of 75 with an SEM of 4, that tells us that the student's knowledge and skills fall between 71 and 79. For average scores at aggregate levels, the number following "±" is the standard error of the mean.
- Performance level data, which are used for tests with performance levels (also known as proficiency levels). Performance levels provide qualitative measurements of students' proficiency in relation to a particular standard or set of standards. Some aggregate reports include performance distribution bars, as in Figure 109, showing the percentage and number of students who achieved each performance level. These bars are color-coded, with three performance levels being coded red-

yellow-green, four being coded red-yellow-green-blue, and five being coded red-yellow-green-blue-purple.

- Percent proficient, also shown in some aggregate reports (see Figure 109), represents the total percentage of students who achieved proficiency. It typically includes those who fell into the top one to three performance levels.

Figure 109. School-Level User View: School Performance on Test Report: Performance by Roster Tab



- Measures in aggregate reports for adaptive tests may also include **Weak or Strong?**, **Proficient?**, and **% Correct**.

In a report, click the more information button in the score or Performance Distribution columns.

A legend appears (see Figure 110 and Figure 111), explaining what the scores or performance levels indicate.

Figure 110. My Students' Performance on Test Report with Expanded Scale Score Legend

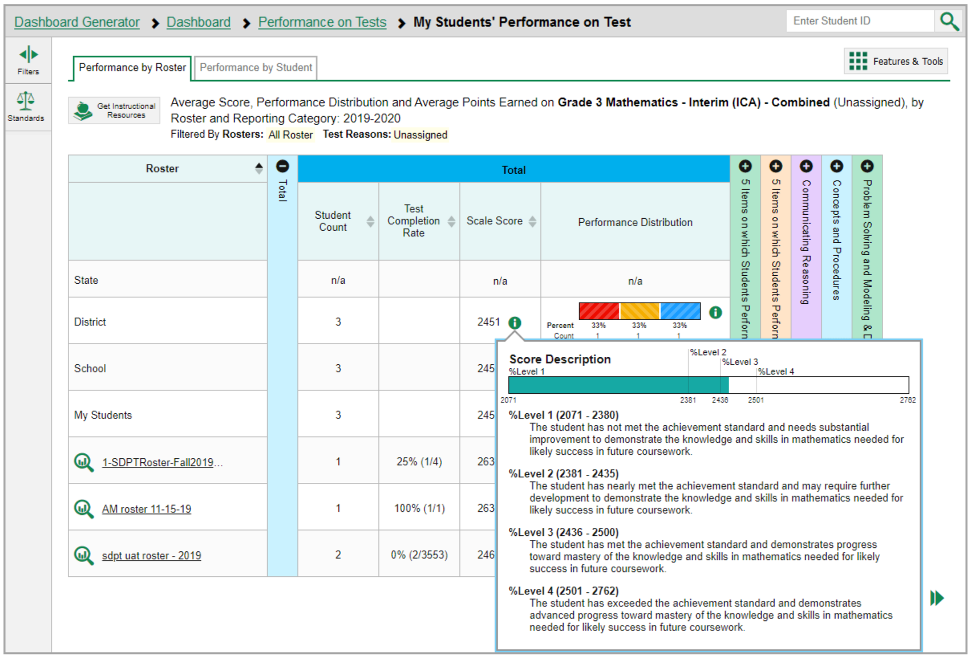
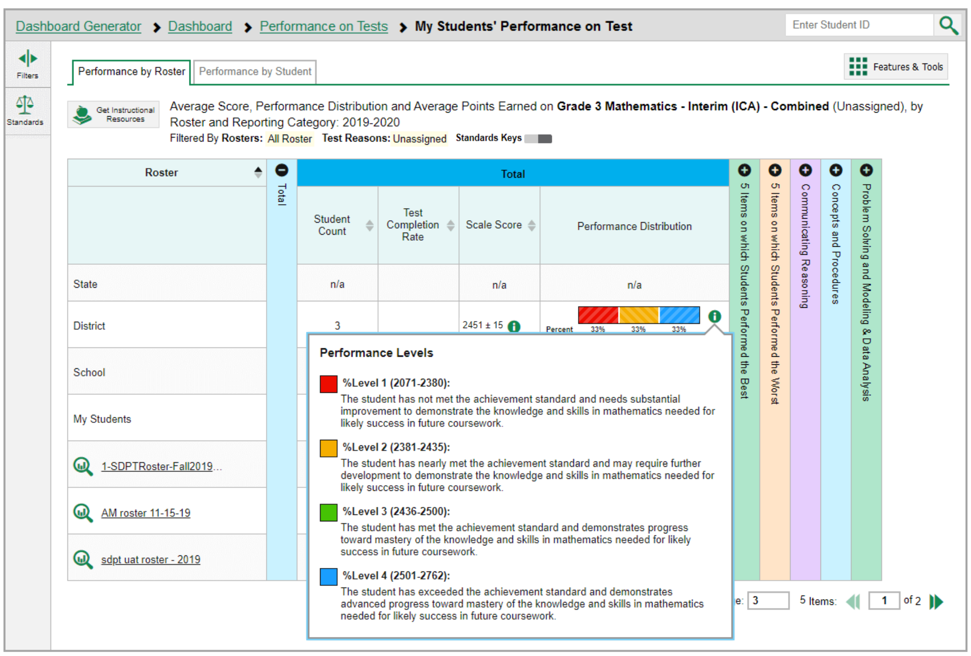


Figure 111. My Students' Performance on Test Report with Expanded Performance Distribution Legend



You will find similar buttons in reports throughout the Reporting System.

R

Report Tables

How to Sort a Table

1. To sort by descending order, click the header of the column you wish to sort by. The bottom arrow in the header is shaded darker when the column is sorted in descending order.

2. To sort by ascending order, click the column header again. The top arrow in the header is shaded darker when the column is sorted in ascending order.

How to Specify the Number of Rows Displayed

In the *Rows per page* field below a table, enter the number of rows you want the table to display per page. Your specifications persist for each table.

You can click the arrow buttons in this field to increase or decrease the number of rows displayed in increments of one.

How to View Additional Table Rows

- To move to the next and previous pages in a table, click the arrow buttons at the lower-right corner of the table.

- To jump to a specific table page, enter the page number in the field 1 at the lower-right corner of the table.

How to View Additional Table Columns

To scroll the table to the right or left, click the arrow buttons on the right and left sides of the table. Alternatively, click and drag the green horizontal scrollbar at the bottom of the table.

If a table contains expandable and collapsible accordion sections, you can click the section bars or and to expand and collapse them.

How to Expand All Accordion Sections in a Table

If you're navigating the page by tabbing through it, you may want to expand all the expandable accordion sections of a table at once. This feature, which is available in most test results, will make the table accessible to a screen reader.

1. Navigate to the table by tabbing through the page in your browser. When the "Load Accessible Table" message appears, press the **Enter** key. All the accordion sections expand.

2. *Optional*: To collapse the sections again, navigate back to the table. When the "Hide Accessible Table" message appears, press the **Enter** key. All the accordion sections collapse, except the **Total** section.

T

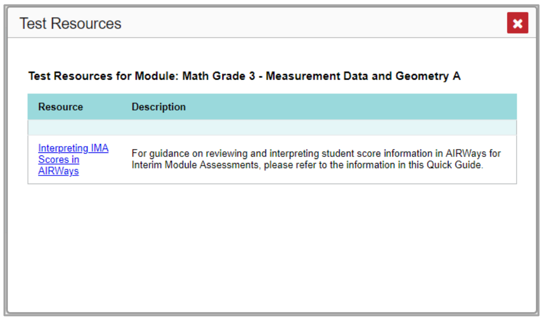
Test Resources

Some test results in the Reporting System include supplementary information that you can access, such as resources provided for the assessment in Tools for Teachers.

If additional assessment information is available, click **Get Instructional Resources** in the **Features & Tools menu**.

A window opens (see Figure 112), displaying resource links that either download or open in a new browser tab or window.

Figure 112. Test Resources Window



U

User Role Change

If your account is associated with multiple user roles or institutions, you can switch between them. The following instructions do not apply to switching schools for hand scoring purposes.

1. From the **Features & Tools** menu, select **Change Role**. (Alternatively, click the **Select Role** link in the path at the upper-left corner.) The ***Change Role*** window appears.
2. From the **Role** drop-down list, select the desired role and institution (entity name) combination.
3. Click **Continue**. The window closes and the Dashboard Generator page appears.

User Support

For additional information and assistance in using the Reporting System, contact the Montana Assessment Helpdesk.

The Helpdesk is open 6:00 AM - 6:00 PM MST (except holidays or as otherwise indicated on the Montana Testing portal).

Montana Help Desk

Toll-Free Phone Support: 877-365-7915

Email Support: mthelpdesk@cambiumassessment.com

- Please provide the Help Desk with a detailed description of your problem, as well as the following:
- If the issue pertains to a student, provide the associated district or school for that student. SSID and additional student PII can only be communicated via phone with the help desk**. Do not leave any student identifying information such as a student name, SSID, and/or personal characteristics in a voicemail or email.**
- If the issue pertains to a Test Information Distribution Engine (TIDE) user, provide the user's full name and email address.
- Any error messages and codes that appeared, if applicable.
- Affected test ID and question number, if applicable.
- Operating system and browser information, including version numbers (for example, Windows 11 (21H2) and Firefox 87 or macOS 12.3 and Safari 13).

Appendix 6-B
Sample Printed Individual Student Report

Individual Student Report

Demo, Student

Student ID: 999999999 | Student DOB: 1/1/2011 | Enrolled Grade: 5

Date Taken: 3/23/2023

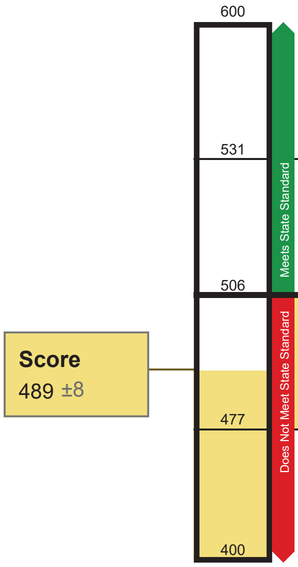
Grade 5 – Science Assessment 2022-2023

Demo District

Demo School

Scale Score: 489±8 **Performance:** Level 2

How Did Your Child Do on the Test?



Level 4 The student demonstrated superior performance in applying more complex science and engineering practices, disciplinary core ideas, and crosscutting concepts, as specified in state standards, to local and global phenomena. The student is becoming a scientifically literate citizen who is on track for content-readiness in science.

Level 3 The student demonstrated solid academic performance to apply science and engineering practices, disciplinary core ideas, and crosscutting concepts, as specified in state standards, to local and global phenomena with application of knowledge to real-world scenarios. The student is on their way to becoming a scientifically literate citizen who is on track for content-readiness in science.

Level 2 The student demonstrated partial mastery to apply science and engineering practices, disciplinary core ideas, and crosscutting concepts, as specified in state standards, to local and global phenomena. The student may require support to become a scientifically literate citizen who is on track for content-readiness in science.

Level 1 The student was, at the time of testing, beginning to demonstrate the ability to apply science and engineering practices, disciplinary core ideas, and crosscutting concepts, as specified in state standards, to local and global phenomena. The student may require significant support to become a scientifically literate citizen who is on track for content-readiness in science.

How Does Your Child's Score Compare?

Name	Average Scale Score
Demo District	498±2
Demo District	501±7
Demo School	458±11

Information on Standard Error of Measurement

A student's score is best interpreted when recognizing that the student's knowledge and skills fall within a score range and not just a precise number. For example, 2300 (±10) indicates a score range between 2290 and 2310.

How Did Your Child Perform on Different Areas of the Test?

The table and the graph below indicate student performance on individual reporting categories. The black dot indicates the student's score on each reporting category. The lines to the left and right of the dot show the range of likely scores your student would receive if he or she took the test multiple times.

Below Standard | At/Near Standard | Above Standard

Category	Performance	Performance Level	Performance level Description
Earth and Space Science	Below the Standard Above the Standard	Below Standard	The student may have trouble integrating content area ideas while investigating the composition, history, and processes that shape earth, the solar system, and the universe.
Life Science	Below the Standard Above the Standard	At/Near Standard	The student can integrate content area ideas while investigating the characteristics, structures, and functions of living things; the processes and diversity of life; and how living organisms interact with each other and their environments.
Physical Science	Below the Standard Above the Standard	Above Standard	The student can integrate content area ideas while investigating how matter and energy exist in a variety of forms and how physical and chemical interactions change matter and energy.

LastName, Test

Student ID: MT9311378 | Student DOB: 1/1/2011 | Enrolled Grade: 8

Date Taken: 4/5/2023

Grade 8 – Science Assessment 2022-2023

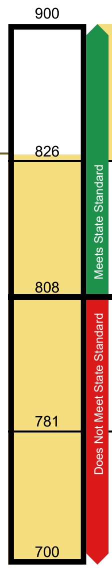
Demo district 9999

Demo Institution 9999_9999

Scale Score: 830±8 **Performance:** Level 4

How Did Your Child Do on the Test?

Score
830 ±8



Level 4 The student demonstrated superior performance in applying more complex science and engineering practices, disciplinary core ideas, and crosscutting concepts, as specified in state standards, to local and global phenomena. The student is becoming a scientifically literate citizen who is on track for content-readiness in science.

Level 3 The student demonstrated solid academic performance to apply science and engineering practices, disciplinary core ideas, and crosscutting concepts, as specified in state standards, to local and global phenomena with application of knowledge to real-world scenarios. The student is on their way to becoming a scientifically literate citizen who is on track for content-readiness in science.

Level 2 The student demonstrated partial mastery to apply science and engineering practices, disciplinary core ideas, and crosscutting concepts, as specified in state standards, to local and global phenomena. The student may require support to become a scientifically literate citizen who is on track for content-readiness in science.

Level 1 The student was, at the time of testing, beginning to demonstrate the ability to apply science and engineering practices, disciplinary core ideas, and crosscutting concepts, as specified in state standards, to local and global phenomena. The student may require significant support to become a scientifically literate citizen who is on track for content-readiness in science.

How Does Your Child's Score Compare?

Name	Average Scale Score
Montana	801
Demo district 9999	754±2
Demo Institution 9999_9999	805±2

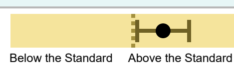
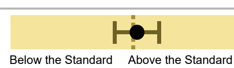
Information on Standard Error of Measurement

A student's score is best interpreted when recognizing that the student's knowledge and skills fall within a score range and not just a precise number. For example, 2300 (±10) indicates a score range between 2290 and 2310.

How Did Your Child Perform on Different Areas of the Test?

The table and the graph below indicate student performance on individual reporting categories. The black dot indicates the student's score on each reporting category. The lines to the left and right of the dot show the range of likely scores your student would receive if he or she took the test multiple times.

Below Standard | At/Near Standard | Above Standard

Category	Performance	Performance Level	Performance level Description
Earth and Space Science		Above Standard	The student can integrate, with mastery, content area ideas while investigating the composition, history, and processes that shape earth, the solar system, and the universe.
Life Science		Above Standard	The student can integrate, with mastery, content area ideas while investigating the characteristics, structures, and functions of living things; the processes and diversity of life; and how living organisms interact with each other and their environments.
Physical Science		At/Near Standard	The student can integrate content area ideas while investigating how matter and energy exist in a variety of forms and how physical and chemical interactions change matter and energy.

