Utah State Assessments 2016–2017 Technical Report

Volume 3:
Test Administration

TABLE OF CONTENTS

LIST OF TABLES

1. OVERVIEW

The set of reports provides comprehensive documentation of the test development, technical quality, test administration procedures, and reporting of test results for Utah's online Student Assessment of Growth and Excellence (SAGE):

Volume 1: 2016–2017 Annual Technical Report

Volume 2: Test Development

Volume 3: Test Administration

Volume 4: Validity

Volume 5: Score Interpretation Guide

Volume 6: Standard Setting

The technical reports describe and summarize the test development, test administration, and statistical and psychometric analyses that were performed on the SY2016–2017 administration of Utah's SAGE English language arts (ELA) assessments in grades 3–11, for mathematics in grades 3–8, end-of-course (EOC) assessments for high school students taking Secondary Mathematics I–III, and for science in grades 4–8, along with end-of-course assessments for high school students taking Earth Science, Biology, Chemistry, and Physics.

Volume 3 describes test administration, test administrator training, security procedures, and test settings and accommodations for the 2016–2017 administration and will be updated whenever significant changes to the administrative procedures included in the assessment system occur.

2. UTAH STATE ASSESSMENTS

The purposes of SAGE were to (1) meet or exceed the requirements of the No Child Left Behind Act (NCLB; 2001) and Utah State Legislature House Bill 15 of 2012, which required district and charter schools to administer computer adaptive tests aligned with Utah Core Standards no later than the 2014–2015 school year; (2) promote and measure the attainment of the Utah Core Standards; and (3) provide information to stakeholders about the assessment, assessment tools, and reports to support teaching and learning.

The Utah state reading, mathematics, and science assessments were required components of the statewide student assessment program. Students' scores in reading, mathematics, and science were included in schools' accountability results. The English language arts (ELA) assessments were administered to students in grades 3–11. The mathematics assessments were administered to students in grades 3–8, and Secondary Math I, II, and III courses. The science assessments were administered to students in grades 4–8, Biology, Chemistry, Earth Science, and Physics courses.

2.1   **TESTING OPTIONS**

The first SAGE testing occurred in spring 2014. This was an operational field test allowing the students to take only one summative test and still providing scale scores and proficiency levels and data for accountability. The 2013–2014 SAGE testing window spanned four months during the school year for the online assessment (February 10–June 13, 2014). Trimester schools tested in February, while the main spring summative window opened on April 1, 2014. The spring testing window for writing opened early to have the writing tests completed in time to schedule range finding meetings. The writing pilot window, provided in late winter, was an operational field test administered over six weeks. The spring testing window was made available in April 2015 for science, math, and ELA (the ELA window let students take writing if they moved to the state after the pilot window).

The initial window of the following school years occurred in summer each year. The second testing window occurred in the fall, where students were provided with both the interim and summative assessments. Both provided scale scores and proficiency levels in real time, as well as data for accountability. The window remained open throughout winter, where trimester schools were given the opportunity to test. The spring testing window was made available for science, math, and ELA.

In September 2016, benchmark modules were made available along with the AIR Ways system for reporting. These are fixed form tests for all subjects, grades, and courses using items from the interim bank. They are short tests each focused on a particular strand. Students have unlimited opportunities and each opportunity expires 10 days from when the test was started.

Computer Adaptive Test Administration

- **SAGE ELA, mathematics, and science:** During the testing window, all eligible students had one opportunity for summative assessments and one opportunity for interim assessments in each content area using the web-based SAGE system.

- **SAGE ELA, mathematics, and science in Braille:** The adaptive SAGE ELA, mathematics, and science assessments were available to students who used Braille. These students had one opportunity to take each content area assessment using new technology and administration procedures.

Table 1. Summary of Tests and Testing Options in 2016–2017

Tests/Subjects	Tested Grades	Number of Testing Opportunities for Summative	Number of Testing Opportunities for Interim or Benchmarks*
SAGE ELA (including writing)	3–11	1	2
SAGE Mathematics	3–8, Secondary Math I, II, and III	1	2
SAGE Science	4 and 8, Biology, Chemistry, Earth Science, and Physics	1	2
SAGE Braille – ELA (including writing)	3–11	1	2

Tests/Subjects	Tested Grades	Number of Testing Opportunities for Summative	Number of Testing Opportunities for Interim or Benchmarks*
SAGE Braille – Mathematics	3–8, Secondary Math I, II, and III	1	2
SAGE Braille – Science	4 and 8, Biology, Chemistry, Earth Science, and Physics	1	2
SAGE Benchmarks – ELA (including writing)	3–11		999
SAGE Benchmarks – Mathematics	3–8, Secondary Math I, II, and III		999
SAGE Benchmarks – Science	4 and 8, Biology, Chemistry, Earth Science, and Physics		999
SAGE Braille Benchmarks – ELA (including writing)	3–11		999
SAGE Braille Benchmarks – Mathematics	3–8, Secondary Math I, II, and III		999
SAGE Braille Benchmarks – Science	4 and 8, Biology, Chemistry, Earth Science, and Physics		999

* Students had two opportunities to take the interim tests—one in the fall window and one in the winter window. Students have unlimited opportunities to take the benchmark assessments.

3. TEST ADMINISTRATION

The SAGE assessments were administered online.

To ensure standardized administration conditions, Test Administrators (TAs) follow procedures outlined in the *Test Administration Manual* (*TAM*). TAs were provided with a specific *TAM* for interim assessments, mid-year summative assessments, and the summative assessment. TAs were urged to review the *TAM* prior to the beginning of testing to ensure that the testing room was prepared for testing, to confirm procedures and awareness of policies of testing prior to students taking the test, and to guarantee knowledge of testing policies.

TAs were required to follow administration procedures and directions. TAs referenced the *TAM* prior to and during testing, ensuring standardized administration conditions for all assessments.

3.1 ADMINISTRATIVE ROLES

The administration of SAGE summative tests required involvement of multiple individuals at each testing site, representing four different roles: School Testing Coordinator (TC), Technology Coordinator, Lab/Session Manager, and Test Administrator (TA)/Proctor. Depending on local policy, a single individual could engage in multiple roles if qualified for each (e.g., the Test Administrator/Proctor may also act as the Lab/Session Manager). These roles and responsibilities are outlined below. For a more detailed description of the roles and responsibilities, please read the *SAGE Test Administration Policies and Procedures.*

3.1.1 School Testing Coordinator

Under the direction of the Assessment Director (AD), the School Testing Coordinator (TC) oversaw all aspects of testing. The TC ensured that Test Administrators carried out the required policies and procedures for standardized testing and that these administrators were properly trained and certified. The TC responsibilities also included working with the ADs to confirm that all teachers and students were properly registered to test with accurate data. Further responsibilities included using the systems to mark special codes and accommodations for appropriate students, verifying proper testing assignments for students, and working with all necessary personnel to resolve testing issues.

3.1.2 Technology Coordinator

The primary responsibility of the Technology Coordinator was to ensure that the school's hardware and software met the requirements for the online assessments. The Technology Coordinator was expected to understand the basic functionality of SAGE, install the secure browser for online testing on each computer prior to testing, and work with the testing coordinators and administrators to coordinate the technical details for testing. For more details on the secure browser used for testing and other hardware and software requirements, please refer to the *SAGE Technical Specifications Manual*.

3.1.3 Lab/Session Manager

The role of a Lab/Session Manager was to work with the Testing Coordinator or Test Administrator/Proctor to distribute student login information. Lab/Session Managers were also responsible for starting, stopping, and pausing all SAGE summative testing sessions, as well as approving students for entry into the testing event using the test delivery system (TDS) and documenting any situations that affected testing (e.g., fire drills, technical issues). The Lab/Session Manager responsibilities, which could be found in the *Test Administration Policies and Procedures*, gave precise documentation of the requirements prior, during, and after testing.

3.1.4 Test Administrator/Proctor

The Test Administrators/Proctors administered the assessments to the students. TAs were expected to ensure that students were able to access the testing session, to actively proctor testing, and to work with the Lab/Session Manager to document and resolve any problems during testing. TAs were also responsible for reviewing the appropriate manuals and user guides on how to administer the assessments, as well as for reviewing the participation report in the Online Reporting System (ORS) with the Testing Coordinator (TC). Finally, TAs were to ensure that all students who required a scheduled make-up test session or completion session were able to do so in order for all students to finish testing. The Test Administrator/Proctor responsibilities, which could be found in the *Test Administration Policies and Procedures*, gave precise documentation of the requirements prior, during, and after testing.

3.2   **ONLINE ADMINISTRATION**

SAGE testing allowed schools to choose testing dates within specified testing windows, to test students in intervals rather than in one long period of time, and to administer the assessments to students for each content area.

To start a test session, the TA would first enter the TA Interface of the online testing system using his or her own computer. A session ID was generated when the test session was opened. Students who were taking the assessment with the TA needed to enter their State Student Identification Number (SSID), first name, and the session ID into the Student Interface using computers provided by the school. New to SY2014–2015 included a change to the structure of SSID numbers. The SSID was shortened to seven digits since the students were no longer required to enter the three leading zeros. In addition, all SAGE eligibility was changed to be controlled by core codes provided by the Utah State Board of Education (USBE) in a nightly UTREx file. The TA then verified that the students were taking the appropriate content area assessment(s) and were provided with the appropriate assessment accommodations, such as use of a Descriptive Audio (see Volume 1 of the Technical Report for a list of accommodations). Students would begin testing only after the TA confirmed that the students were taking the appropriate assessments(s) and approved them to be tested. The TA would then read the *Directions for Administration* in the SAGE *TAM* aloud to the students and walk them through the login process.

Once an assessment was started, students had to answer all test questions before proceeding to the next question; students were not allowed to skip questions. Students were permitted to select items to review at the end of the test. The online testing system allowed a student to scroll back to review and edit answers, as long as the student was in the same test session and the test session had not been paused for more than 20 minutes. The pause rule was not enforced on the writing test. In the online testing system, an assessment could be started in one test session and completed in another session(s). In a subsequent test session, answers provided in the previous test session would not be available for review or editing if the time between sessions were more than the pause rule allowed, except for writing assessments. Test sessions were not timed; therefore, students could use as much time as needed to complete an assessment.

TAs could also pause a single student's assessment, or all the assessments during a test session (for example, to give students a break). It was up to the TA to determine an appropriate stopping point; however, assessments were not paused for more than 20 minutes to ensure the integrity of the assessments with the exception of the writing test.

The TA should have remained in the room at all times during a test session to monitor student testing. Once the test session ended, the TA would make sure that each student had successfully logged out of the system, collect any handouts or scratch paper that was used by students during the assessment, and securely shred them.

3.3   **BRAILLE AND AMERICAN SIGN LANGUAGE TEST ADMINISTRATION**

In SY2014–2015, SAGE was made available to students who use Braille as a mode of instruction, allowing these students to have access to the adaptive SAGE assessments.

The Braille interface of SAGE delivers assessments to students in the following formats:

- The Braille interface worked with the Job Access with Speech (JAWS) Screen Reading software provided by Freedom Scientific and was an essential component that students used with the Braille interface.

- SAGE mathematics and science items were presented to students in Nemeth Braille through the adaptive SAGE test via a Braille embosser.

- Students taking the SAGE ELA were able to emboss both reading passages and items as they progressed through the assessment. If a student had Refreshable Braille Display (RBD), a 40-cell RBD was recommended. The SAGE ELA was presented to the student with items in contracted Literary Braille (for items containing only text) and via a Braille embosser (for items with tactile or spatial components that could not be read by a Refreshable Braille Display).

Prior to administering SAGE assessments using the Braille interface, TAs were required to ensure that the technical requirements were met. These requirements applied to the student's computer, the TA's computer, and the supporting Braille technologies used in conjunction with the Braille interface. Any additional requirements were outlined in each of the respective *TAM*s.

USBE has made the decision to transition to UEB braille for all subjects. The Department is allowing students IEPs to determine whether they will receive EBAE/Nemeth or UEB braille. In the spring of 2017, one student required UEB braille for his/her grade 5 summative math test. AIR had enough grade 5 math items brailled to meet blueprint and delivered this test.

Starting with the spring 2015 administration, the American Sign Language (ASL)–embedded accommodation became available for ELA listening stimuli. Using this accommodation, students were able to retrieve ASL videos to help access test content. The videos appeared in a window on the same screen as the items, showing a human signer translating test content.

3.4 ALLOWABLE RESOURCES FOR ONLINE TESTING

During testing, students could use specified tools and resources, including scratch and graph paper, pencils, or pens. A pop-up reference sheet was available on the Chemistry and Physics tests. TAs could print out the reference sheet for students as well. Table 2 provides resources that may be available to students during assessments.

Table 2. Allowable Resources for SAGE for SY2016–2017

Test	Allowed Materials
SAGE Summative and Interim Assessments	Headphones Scratch and/or graph paper Pencil and/or pen
All Science Grade 6, 7, 8 and Secondary Math I, II, and III	Any non-Internet-capable calculator the student used during instruction (an onscreen calculator will also be available)
Physics	Physics reference sheet (also embedded in test)
Chemistry	Chemistry reference sheet (also embedded in test)

4. TRAINING AND INFORMATION FOR TEST COORDINATORS AND ADMINISTRATORS

Test Coordinators (TCs) oversaw all aspects of testing at their schools and served as the main point of contact, while TAs administered the online assessments. Webinars, user guides, manuals, and training sites were used to train TAs about the online testing requirements and the mechanics of starting, pausing, and ending a test session. Training materials for the administration were provided online.

In addition, literature was provided on helping students, their families, teachers, and districts to understand their scores that were provided starting in October 2014. This included brochures, interpretative guides in English and Spanish, and an explanatory video and PowerPoint detailing the different aspects of student scores and the individual student reports.

4.1 ONLINE TRAINING

Multiple training opportunities were offered online to key staff throughout the year, which included webinars and training tests.

4.1.1 Webinars

USBE, with AIR support, offered both in-person and webinar presentations leading up to the administration for a variety of users. The first part of the trainings focused on the technology enhancements and how to access the online testing system. The second part of the trainings covered scheduling and conducting test sessions. The webinars were primarily PowerPoint presentations.

The interactive nature of these training webinars allowed the participants to ask questions during and after the presentation. There were practice activities that followed the trainings and allowed future users to get a hands-on experience with the systems. These dates and locations are listed below.

- August 26, 2016: New Assessment Director meeting
- August 31, 2016: Fall Training Webinar from USBE
- September 1, 2016: In-person SAGE training for New Assessment Directors
- January 31, 2017: Winter Training Session at the Jordan School District
- February 1, 2017: Winter Training Webinar from USBE

4.1.2 Training Sites

The SAGE training test site was available 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 an online assessment on the student training site. The student training test site contained approximately 15–20 test items per grade band and content area. A student could log in directly to the training site as a "Guest" without a TA-generated test session ID or could log in through a training test session created by the TA in the TA training site. Items in the student training test included all item types that were included in the operational item pool (i.e., multiple-choice items, grid items, and natural language items).

The training test was also equipped with the same tools provided in the summative and interim tests, including the dictionary tool. Updated tutorials were also provided by December 2014. Students with hearing impairments had the option of an ASL Video setting whereby they could watch a signed video of the listening stimulus on the training test. Braille items were also made available to practice in the training test.

In September 2016, practice tests for math and ELA were added to the training site. The practice tests were created from released Independent College and Career Readiness (ICCR) bank items. They are fixed form tests about half the length of the summative tests and representative of the interim blueprint. Upon submitting a practice test, the student would see an interactive report showing his or her score on each item and giving the student the opportunity to return to the item to see and change his or her response.

4.2 MANUALS AND USER GUIDES

In addition to the online training and resources, a series of manuals and user guides were available on the SAGE portal (www.sageportal.org).

The *SAGE Technical Specifications Manual for Online Testing* and the *SAGE System Requirements* documents were available online and provided both information and resources for the Technology Coordinator and Test Administrator roles. They covered the hardware and the software requirements for SAGE and information about the secure browsers. The *SAGE Technical Specifications Manual for Online Testing* provided information about supported operating systems and related requirements, network and Internet requirements, general hardware and software requirements, and text-to-speech information. Instructions for specific software configuration changes were also described in the manual.

The *SAGE Test Information Distribution Engine (TIDE) User Guide* was available online and provided information about the TIDE application within the SAGE system. This application allowed users to manage user role assignments, set student accommodations for testing, and update user information.

The *SAGE Test Administration Manual* served as a software guide on how to use the online system applications, including the TA and student testing sites. The *SAGE Test Administration Manual* outlines the roles and responsibilities of individuals involved in administering the SAGE summative tests at each testing site, representing four different roles: School Testing Coordinator (TC), Technology Coordinator, Lab/Session Manager, and Test Administrator (TA)/Proctor. Depending on local policy, a single individual could engage in multiple roles if qualified for each (e.g., the Test Administrator/Proctor may also act as the Lab/Session Manager). The Summative and Interim *SAGE Test Administration Manuals* are presented as Appendix A and B, respectively.

The *SAGE Online Reporting User Guide* provided instructions on how to generate reports to see which students have or have not completed assessments.

The *SAGE Secure Browser Installation Manual* was developed to provide instructions for installing secure browsers on computers and devices running a supported operating system.

The *SAGE Braille Requirements and Testing Manual* provided information about supported hardware and software requirements and how to configure JAWS. The manual shared general

information about administering a test to a student with a Braille accommodation; printing test material was also included. A separate frequently asked questions document was later provided during the testing window.

The *SAGE Special Codes User Guide* was created to address the use of special codes for accountability purposes.

The *SAGE Calculator Manual* was created to provide steps for schools to access the Desmos calculators used in SAGE throughout the year.

The *End of Support Plan* gave clear guidelines regarding which operating systems AIR would support for the upcoming test administration and in following years. The purpose of this document was to help Local Education Agencies (LEAs) and schools manage support requirements and allow for both product and information technology planning in organizations based on knowledge of the support timelines.

The *Utah Accommodations and Participation Policy* was created by USBE to address the Department's policy on student participation and accommodations.

All manuals and user guides were available on the SAGE portal prior and during the testing window.

5. TEST SECURITY

This section describes test security, student confidentiality, and policies on testing impropriety. The SAGE assessment system incorporates security systems and procedures across the range of test activities, from item and test development through test administration, scoring, and reporting. Secure test systems prevent unauthorized access to confidential student information and test content, real-time forensic analysis reports monitor testing to detect irregularities, and extensive training reinforces standardized test administration procedures, including procedures to report and document violations. These systems and procedures are consistent with best practices described in the *TILSA Test Security Guidebook* (Olson and Fremer, 2013; see also Wollack and Fremer, 2013). SAGE test security procedures are also detailed in Volume 1 Section 11.

5.1 STUDENT-LEVEL TESTING CONFIDENTIALITY

The Family Educational Rights and Privacy Act (FERPA) prohibits the public disclosure of student information or test results. The following are examples of prohibited practices:

- Giving out login information (username and password) either to other authorized TIDE users or to unauthorized individuals.

- Sending a student's name and SSID number together in an e-mail message. If information must be sent via e-mail or fax, include only the SSID number, not the student's name.

- Having students log in and test under another student's SSID number.

Student test materials and reports could not be exposed in such a manner that student names could be identified with student results, except by authorized individuals with an educational need to know.

All students, including home-schooled students, were required to be enrolled or registered at their testing schools in order to take the online or Braille assessments. Student enrollment information, including demographic data, is uploaded to the SAGE systems from the UTREx system nightly via a secured file transfer site to the online testing system.

Students logged in to SAGE using their legal first name, SSID number, and a test session ID. Only students could log in to an online test session. TAs, proctors, or other personnel were not permitted to log in to the SAGE system on behalf of students, although they were permitted to assist students who needed help logging in.

5.2   TEST SECURITY

The importance of maintaining test security and the integrity of test items was stressed throughout the webinar trainings and in the user guides and manuals. Features in the testing system also protected test security.

5.2.1   System Built-in Test Security

- A hierarchy of control: Lab/Session Managers, Technology Coordinators, Testing Coordinators, and Test Administrators had well-defined roles and access to the testing system. USBE provided the list of active LEA administrators. These LEA administrators were responsible for managing all other users in their LEA. Throughout the year, the LEAs were also expected to delete information in TIDE for any staff members who had transferred to other schools, resigned, or no longer served as TAs or teachers.

- Password protection: All access points by different roles—at the state level, LEA level, school level, and school staff level—required a password to log in to the system. Newly added TAs and teachers received separate passwords through their personal e-mail addresses assigned by the school. Additional password requirements were created to increase the strength of user passwords. These requirements included that passwords have a minimum of eight characters and include an uppercase letter, lowercase letter, a number, and a symbol.

- Secure browser: A key role of the Technology Coordinator was 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 prevented students from accessing other computers or Internet applications and from copying test information. The secure browser suppressed access to commonly used browsers such as Internet Explorer and Firefox, and prevented students from searching for answers on the Internet or communicating with other students. SAGE could be accessed only through the secure browser and not by other Internet browsers.

5.2.2   Test Security and Ethics

SAGE summative and interim tests are highly secure and should be treated as such. Access to the SAGE summative and interim testing systems could only be provided to qualified personnel. Because students used the same personal information for each test they took, proctors could only allow access to tests for students who were physically present in the room with them. No access

to secure test materials could be granted to anyone who was not a student scheduled to take an exam. Non-students could not access test content at any time.

All test materials could be handled only by qualified personnel, and a system of materials accounting could be in place to ensure that all test materials are accounted for at the conclusion of testing. TAs securely stored all **used and unused** test materials. Students were not allowed to remove test content from a testing session. Students could not store test content or questions on their calculators. All writing done by students on scratch paper, graph paper, or formula sheets must have been destroyed at the conclusion of the testing window.

Educators could not examine test content, including passages, questions, or answer options, at any time. Under no circumstances could actual passages, prompts, or questions from these tests be taught to or reviewed with students.

The validity of the test was compromised when students received assistance on the test either explicitly by prior knowledge of questions, or implicitly by modified instruction by the educator after test content has been examined.

Reproducing the test via electronic or paper means was not permitted. Such practices violated test security and testing ethics. According to state law, evidence of these illegal activities could result in disciplinary action and/or the loss of teacher licensure.

Educators could not read passages, questions, or answer options to a student. All students have access to the text-to-speech tool throughout each test, with the exception of the reading passages in the ELA tests.

For further information regarding testing ethics and test security, see the Utah State Board of Education–approved *Standard Test Administration and Testing Ethics Policy*.

The new Benchmark tests are reported using AIR Ways and allows teachers to view the items and student responses.

Scheduling Make-Up Testing and Test Completion Sessions

Test completion sessions could include students working on different tests.

Unexpected circumstances (e.g., fire drills, power failures) could interrupt testing. Test completion sessions could be scheduled when normal conditions were restored. Interruptions could not reduce the total amount of time students were given to complete tests.

After a test had been paused for 20 minutes, the student could no longer view or modify responses from that testing session. Students could not view or change prior answers during a make-up session. A make-up or completion session was only to finish the remaining portions of the test. This limit did not apply to the ELA writing test, which could be modified up to the point of submission.

Test Irregularities

On rare occasions, a non-standard situation arose during test administration. Three ways to account for irregularities were provided. Steps for dealing with test irregularities are outlined in more detail in Section VI of the *TIDE User Guide*.

Reset a Test

Resetting a test eliminates all responses for a student. When that student logged in to the test again, the test would start over. This could only be done in situations where the test could not be appropriately completed as is (e.g., two students accidentally log in to each other's test, a student requiring Braille was not given the accommodation, etc.). A test could never be reset to give a student a second opportunity.

Grace Period Extension

Extending the grace period of a test gives a student access to all of his or her previous responses. This could be granted if a test session is unexpectedly interrupted (e.g., fire drill, lockdown, etc.). The grace period extension could not be applied if the test session ended normally or if the student was given time to review his or her answers before logging out of a test.

Invalidate a Test

Tests could be invalidated when a student's performance was not an accurate measure of his or her ability (e.g., the student cheated, used inappropriate materials, etc.). If a test was invalidated, the student was not given another opportunity to take the test. Invalidating a test required the approval of an LEA-level user.

Reopen a Test

Reopening a test changed the test's status from completed or reported to paused. This capability was useful if a student accidently submitted a test prior to reviewing it. After the test was reopened, a student could resume testing. A test was not reopened once a student saw a score.

Reopen a Test Segment

Reopening a test segment will allow a student to return to a prior segment in cases where the student moved to the next segment in error. This could occur on both summative and interim math grade 6 tests or summative writing tests. After the test segment was reopened, a student could return to the prior segment and complete his or her work.

5.3 ONLINE MANAGEMENT SYSTEM

We employ various measures described in the proposal to ensure data are secured from breaches and identity theft through implementation of physical, network, and software security protections. Beyond breaches and theft, all of our secure websites and software systems enforce role-based security models that protect individual privacy and confidentiality in a manner consistent with Utah's privacy laws, the Family Educational Rights and Privacy Act (FERPA), and other federal laws. Our systems implement sophisticated, configurable privacy rules that can limit access to data to only appropriately authorized personnel. Different states interpret FERPA differently, and we support customized interpretations. Our systems are designed to support these interpretations flexibly. AIR is committed to working with the Department to maintain data security according to its specifications.

With regard to the Children's Online Privacy Protection Act (COPPA), AIR is not collecting any personal information directly from children and, as such, does not have procedures in place to

obtain parental consent. We assume USBE is covered by statutes authorizing the use of such data in the student assessment program. However we receive the data, AIR will follow all COPPA requirements to maintain the confidentiality, security, and integrity of personal information we receive; retain such information collected for only as long as is necessary to support testing and reporting; and delete the information using reasonable measures to protect against its unauthorized access or use.

The Federal Information Security Management Act (FISMA) addresses the confidentiality, integrity, and availability of data in federal agencies and federal contractors and does not appear to be directly applicable to the Utah data under this contract. However, as part of FISMA, the National Institute of Standards and Technology is responsible for guidance and standards, including minimum requirements, for providing adequate information security.

5.3.1 Secure System Design

AIR has developed a custom single sign-on application that is made available in Utah's secure portal. This application is used to support access to AIR's system in accordance with Utah's user ID and password policy. Authorized users can log in to Utah's single sign-on using their current user IDs and passwords and can be redirected to AIR's portal, where they have access to AIR's secure applications such as the Test Information Distribution Engine (TIDE), the test delivery system (TDS), and Online Reporting System (ORS). Nightly backups protect the data. The server backup agents send alerts to notify system administration staff in the event of a backup error, at which time they will inspect the error to determine whether the backup was successful or they will need to rerun the backup. The system can withstand failure of almost any component with little or no interruption of service.

AIR's hosting provider, Rackspace, has redundant power generators that can continue to operate for up to 60 hours without refueling. With the multiple refueling contracts that are in place, these generators can operate indefinitely. Rackspace partners with nine different network providers, providing multiple, redundant data routes. Every installation is served by multiple servers, any one of which can take over for an individual test upon failure of another.

AIR's architecture ensures data are recoverable at all times. Each disk array is internally redundant, with multiple disks containing each data element. Immediate recovery from failure of any individual disk is performed by accessing the redundant data on another disk. AIR maintains support and maintenance agreements through our hosting provider for all of the hardware used by our systems.

5.3.2 System Security Components

AIR has built-in security controls in all of its data stores and transmissions. Unique user identification is a requirement for all systems and interfaces. All of AIR's systems encrypt data at rest and in transit.

Physical Security

USBE data reside on servers at Rackspace, AIR's hosting provider. Rackspace maintains 24-hour surveillance of both the interior and exterior of its facilities. All access is keycard controlled, and sensitive areas require biometric scanning.

Secure data are processed at AIR facilities and are accessed from AIR machines. AIR's servers are in a secure, climate-controlled location with access codes required for entry. Access to our servers is limited to our network engineers, all of whom, like all AIR employees, have undergone rigorous background checks.

Staff at both AIR and Rackspace receive formal training in security procedures to ensure that they know the procedures and implement them properly. AIR and Rackspace protect data from accidental loss through redundant storage, backup procedures, and secure off-site storage.

Network Security

Hardware firewalls and intrusion detection systems protect our networks from intrusion. They are installed and configured to prevent access for services other than hypertext transfer protocol secure (HTTPS) for our secure sites.

AIR's systems maintain security and access logs that are regularly audited for login failures, which may indicate intrusion attempts.

Software Security

All of AIR's secure websites and software systems enforce role-based security models that protect individual privacy and confidentiality in a manner consistent with Utah's privacy laws, the Family Educational Rights and Privacy Act (FERPA), and other federal laws.

AIR's systems implement sophisticated, configurable privacy rules that can limit access to data to only appropriately authorized personnel. Different states interpret the FERPA differently, and our system is designed to support these interpretations flexibly. AIR has worked with USBE to maintain data security according to its specifications.

AIR maintains logs of key activities and indicators, including data backup, server response time, user accounts, system events and security, and load test results. In addition, AIR runs automated functional tests of our test delivery system every morning, and logs from these runs are available for at least one week from the time of the run.

AIR psychometricians monitor the quality and performance of test administrations statewide through a series of quality assurance (QA) reports. The QA reports provide information on item behavior, blueprint match rates, and item exposure rates, and also provide cheating analysis reports.

5.3.3 Quality Control for Data, Analyses, Scoring, and Score Reports

AIR implements a series of quality control steps to ensure error-free production of score reports in an online format. The quality of the information produced in the test delivery system (TDS) is tested thoroughly before, during, and after the testing window.

5.4 BEFORE TESTING WINDOW

Prior to its implementation in the operational test administration, the AIR scoring engine and the accuracy of data files are checked, using a simulated student response data file. The simulated data are used to check whether the student responses entered in the TDS were captured accurately and the scoring specifications were applied accurately. The simulated data file is scored independently by two programmers, following the scoring rules.

In addition to checking the scoring accuracy, the test configuration file is checked thoroughly. For the operational administration, a test configuration file is the key file that contains all specifications for the item selection algorithm, and eventually for the scoring algorithm, such as the test blueprint specification, slopes and intercepts for theta-to-scale score transformation, cut scores, and the item information (cut scores, answer keys, item attributes, item parameters, passage information, etc.). The accuracy of the information in the configuration file is checked and confirmed numerous times independently by multiple staff members prior to the testing window.

5.4.1 Web Approval of Content During Development

The Item Tracking System (ITS) integrates directly with the test delivery system (TDS) display module, and displays each item exactly as it will appear to the student. This process is called web preview, and web preview is tied to specific item review levels. Upon approval at those levels, the system locks content as it will be displayed to the student, transforming the item representation to the exact representation that will be rendered to the student. No change to the display content can occur without a subsequent web preview. This process freezes the display code that will present the item to the student.

Web approval functions as an item-by-item blueline review. It is the final rendering of the item as the student will see it. Layout changes can be made after this process in two ways:

1. Content can be revised and re-approved for web display.
2. Online style sheets can change to revise the layout of all items on the test.

Both of these processes are subject to strict change control protocols to ensure that accidental changes are not introduced. Below, we discuss automated quality control processes during content publication that raise warnings if item content has changed after the most recent web-approved content was generated. The web approval process offers the benefit of allowing final layout review much earlier in the process, reducing the work that must be done during the very busy period just before tests go live.

5.4.2   Platform Review

Platform review is a process in which each item is checked to ensure that it is displayed appropriately on each tested platform. A platform is a combination of a hardware device and an operating system. In recent years, the number of platforms has proliferated, and platform review now takes place on approximately 15 platforms that are significantly different from one another.

Platform review is conducted by a team. The team leader projects the item as it was web approved in ITS, and team members, each behind a different platform, look at the same item to see that it renders as expected.

5.4.3   Testing, Deployment to Production, and Maintenance of the Production Environment

Each release of every one of our systems goes through a complete testing cycle, including regression testing. Each release, and every time we publish a test, the system goes through user acceptance testing (UAT). During UAT, we provide our client with login information to an identical (though smaller scale) testing environment to which the system has been deployed. We provide recommended test scenarios and constant support during the UAT period. For Utah, we began UAT a full 4 weeks prior to opening of the testing window. Issues identified within the first 10 calendar days of testing were resolved and the fixes returned to production by the 14th calendar day of testing for final system check.

Deployments to the production environment all follow specific, approved deployment plans. Teams working together execute the deployment plan. Each step in the deployment plan is executed by one team member and verified by a second. Each deployment undergoes shakeout testing following the deployment.

This careful adherence to deployment procedures ensures that the operational system is identical to the system tested on the testing and staging servers. Upon completion of each deployment project, management approves the deployment log.

During the course of the year, some changes may be required to the production system. Outside of routine maintenance, no change is made to the production system without approval of the Production Control Board (PCB). The PCB includes the director of AIR's Assessment Program or the chief operating officer, the director of our Computer and Statistical Sciences Center, and the project director, in this case Robin Seldin. Any request for a change to the production system requires the signature of the system's lead engineer. The PCB reviews risks, test plans, and test results. In the event that any proposed change will affect client functionality or pose risk to operation of a client system, the PCB ensures that the client is informed and in agreement with the decision.

The PCB approves a maintenance plan that includes every scheduled change to the system.

Deviations from the maintenance plan must be approved by the PCB, including server or driver patches that differ from those approved in the maintenance plan.

Every bug fix, enhancement, data correction, or new feature must be presented with the results of a quality assurance plan and approved by the PCB.

An emergency procedure is in place that allows rapid response in the event of a time-critical change needed to avert compromise of the system. Under those circumstances, any member of the PCB can authorize the senior engineer to make a change, with the PCB reviewing the change retroactively.

Typically, deployments happen during a maintenance window, and deployments are scheduled at a time that can accommodate full regression testing on the production machines. Any changes to the database or procedures that in any way might affect performance are typically subject to a load test at this time.

Cutover and Parallel Processing

AIR maintains multiple environments to ensure smooth cutover and parallel processing. With a centralized hosting site in Washington, DC, multiple development environments and a test environment can be maintained. At Rackspace, we maintain a staging environment and the production environment.

The production environment runs independently of the other environments and is only changed with the approval of the PCB. When developing enhancements, they are developed and tested initially on the development and test environments in Washington, DC before being deployed to the staging environment in Rackspace.

The staging environment is a scaled-down version of the production environment. It is in this environment that UAT takes place. Only when UAT is complete and the PCB signs off is the production environment updated. In this way, the system continues to function uninterrupted as testing takes place in parallel until a clean cutover takes place.

5.4.4 Functionality and Configuration

The items, both in themselves and as configured onto the tests, form one type of online product. The delivery of that test can be thought of as an independent service. Here, we document quality assurance procedures for delivering the online assessments.

One area of quality unique to online delivery is the quality of the delivery system. Three activities provide for the predictable, reliable, quality performance of our system:

1. Testing on the system itself to ensure function, performance, and capacity
2. Capacity planning
3. Continuous monitoring

AIR statisticians examine the delivery demands, including the number of tests to be delivered, the length of the window, and the historic state-specific behaviors to model the likely peak loads. Using data from the load tests, these calculations indicate the number of each type of server necessary to provide continuous, responsive service, and AIR contracts for service in excess of this amount. Once deployed, our servers are monitored at the hardware, operating system, and

software platform levels with monitoring software that alerts our engineers at the first signs that trouble may be ahead. Applications log not only errors and exceptions, but latency (timing) information for critical database calls. This information enables us to know instantly whether the system is performing as designed, or if it is starting to slow down or experience a problem.

In addition, latency data is captured for each assessed student—data about how long it takes to load, view, or respond to an item. All of this information is logged as well, enabling us to automatically identify schools or districts experiencing unusual slowdowns, often before they even notice.

5.5 DURING TESTING WINDOW

5.5.1 Quality Assurance in Test Scoring

AIR verifies the accuracy of the scoring engine using simulated test administrations. The simulator generates a sample of students with an ability distribution that matches that of the state. The ability of each simulated student is used to generate a sequence of item responses consistent with the underlying ability. Although the simulations were designed to provide a rigorous test of the adaptive algorithm for adaptively administered tests, they also provide a check of the full range of item responses and test scores in fixed form tests as well. Simulations are always generated using the production item selection and scoring engine to ensure that verification of the scoring engine is based on a very wide range of student response patterns.

To verify the accuracy of the Online Reporting System, we merge item response data with the demographic information taken either from previous year assessment data, or if current year enrollment data is available by the time simulated data files are created, we can verify online reporting using current year testing information. By populating the simulated data files with real school information, it is possible to verify that special school types and special districts are being handled properly in the reporting system.

Specifications for generating simulated data files are included in the analysis specifications document submitted to the Department each year. Review of all simulated data is scheduled to be completed prior to the opening of the test administration, so that the integrity of item administration, data capture, item and test scoring and reporting can be verified before the system goes live.

To monitor the performance of the assessment system during the test administration window, a series of quality assurance reports can be generated at any time during the online assessment window. For example, item analysis reports allow psychometricians to ensure that items are performing as intended and serve as an empirical key check through the operational test window. In the context of adaptive test administrations, other reports such as blueprint match and item exposure reports allow psychometricians to verify that test administrations conform to specifications.

An additional set of cheating analysis reports flags unlikely patterns of behavior in testing administrations. The quality assurance reports are generated on a regular schedule. Item analysis

and blueprint match reports are evaluated frequently at the opening of the test window to ensure that test administrations conform to blueprint and items are performing as anticipated.

Each time the reports are generated, the lead psychometrician reviews the results. If any unexpected results are identified, the lead psychometrician alerts the project manager immediately to resolve any issues. Table 3 presents an overview of the quality assurance (QA) reports.

Table 3. Overview of Quality Assurance Reports

QA Reports	Purpose	Rationale
Item Statistics	To confirm whether items work as expected	Early detection of errors (key errors for selected-response items and scoring errors for constructed-response, performance, or technology items)
Item Exposure Rates	To monitor unlikely high exposure rates of items or passages or unusually low item pool usage (high unused items/passages)	Early detection of any oversight in the blueprint specification
Cheating Analysis	To monitor testing irregularities	Early detection of testing irregularities

Item Analysis Report

The item analysis report is used to monitor the performance of test items throughout the testing window and serves as a key check for the early detection of potential problems with item scoring, including incorrect designation of a keyed response or other scoring errors, as well as potential breaches of test security that may be indicated by changes in the difficulty of test items. To examine test items for changes in performance, this report generates classical item analysis indicators of difficulty and discrimination, including proportion correct and biserial/polyserial correlation, as well as item response theory (IRT)–based item fit statistics. The report is configurable and can be produced so that only items with statistics falling outside a specified range are flagged for reporting or to generate reports based on all items in the pool.

Item p-Value. For multiple-choice items, the proportion of students selecting each response option is computed; for constructed-response, performance, and technology items, the proportion of student responses classified at each score point is computed. For multiple-choice items, if the keyed response is not the modal response, the item is also flagged. Although the correct response is not always the modal response, keyed response options flagged for both low biserial correlations and non-modal response are indicative of miskeyed items.

Item Discrimination. Biserial correlations for the keyed response for selected-response items and polyserial correlations for polytomous constructed response, performance, and technology items are computed. AIR psychometric staff evaluates all items with biserial correlations below a target level, even if the obtained values are consistent with past item performance.

Item Fit. In addition to the item difficulty and item discrimination indices, an item fit index is produced for each item. For each student, a residual between observed and expected score given

the student's ability is computed for each item. The residuals for each are averaged across all students, and the average residual is used to flag an item.

Cheating Analysis Report

The validity of test score interpretation depends critically on the integrity of the test administrations on which those scores are based. Any irregularities in the administration of assessments can therefore cast doubt on the validity of the inferences based on those test scores. Multiple facets ensure that tests are administered properly, including clear test administration policies, effective test administrator training, and tools to identify possible irregularities in test administrations.

For online administrations, quality assurance (QA) reports are generated during and after the test windows. These are geared toward detection of possible cheating, aggregating unusual responses at the student level to detect possible group-level testing anomalies. Compared to fixed form assessments, the adaptive assessments reduce potential opportunities for cheating because students receive varying sets of items.

Online test administration allows information to be tracked that was not possible to track in the context of the paper-and-pencil tests. This information includes not only item responses but also number of revisits to an item or items, test start and end times, scores in each opportunity in the current year, scores in the previous year, and other selected information in the system (e.g., accommodations) as requested by the state. AIR's test delivery system (TDS) captures all of this information.

Unlike with paper assessments where data analysis must await the close of test window and processing of answer documents, AIR's TDS allows AIR psychometricians and state assessment staff to monitor testing anomalies throughout each test administration window, following the first operational administration. Following the base year, the analyses used to detect the testing anomalies can be run anytime within the testing window. Evidence evaluated included changes in test scores across administrations, item response time, and item response patterns using the person-fit index. The flagging criteria used for these analyses are configurable and can be changed by the user. Spring 2016 reports were reviewed following scoring of operational tests.

5.6   DATA PREPARATION AND QUALITY CHECK

When a student responds to test questions online, his or her response to each item is immediately captured and stored in the Database of Record (DoR) at AIR, a repository for all data relevant to a student's testing experience. Our quality assurance procedures are built on two key principles: automation and replication. Certain procedures can be automated, which removes the potential for human error. Procedures that cannot be reasonably automated are replicated by two independent analysts at AIR.

When data are prepared for psychometric analyses, they undergo two phases: a data preparation phase and a psychometric phase. In the former phase, data are extracted from the DoR and provided to two independent SAS programmers. These two programmers are provided with the client-assigned business rules, and they independently prepare data files suitable for subsequent

psychometric analysis. The data files prepared by the different programmers are formally compared for congruency. Any discrepancies identified are resolved through code review meetings with the programmer lead and the lead psychometrician.

When the two data files match exactly, they are then passed over to two independent psychometricians, who each perform classical and IRT analyses. Any discrepancies are identified and resolved.

When all results match from the independent analysts, the final results are uploaded to AIR's Item Tracking System (ITS).

5.7   SCORE REPORT QUALITY CHECK

Scores for the SAGE online assessments are assigned by automated systems in real time. For machine scored portions of assessments, the machine rubrics are created and reviewed along with the items, then validated and finalized during rubric validation following field-testing. The review process "locks down" the item and rubric when the item is approved for web display (Web Approval). During operational testing, actual item responses are compared to expected item responses (given the item response theory [IRT] parameters), which can detect miskeyed items, item drift, or other scoring problems. Potential issues are automatically flagged in reports available to psychometricians.

The handscoring processes for writing tasks include rigorous training, validity and reliability monitoring, and back-reading to ensure accurate scoring. Handscored items are married up with the machine-scored items by our Test Integration System (TIS). The integration is based on identifiers that are never separated from their data and are further checked by the quality monitor (QM) system where the integrated record is passed for scoring. Once the integrated scores are sent to the QM, the records are rescored in the test-scoring system that applies the SAGE scoring rules and assigns scores from the calibrated items, including calculating performance level indicators, subscale scores and other features, which then pass automatically to the reporting system and Database of Record (DoR). The scoring system is tested extensively prior to deployment, including hand checks of scored tests and large-scale simulations to ensure that point estimates and standard errors are correct.

After passing through the series of validation checks in the QM system, data are passed to the DoR, which serves as the centralized location for all student scores and responses, ensuring there is only one place where the "official" record is stored. Only after scores have passed the QM checks and are uploaded to the DoR are they passed to the Online Reporting System, which is responsible for presenting individual-level results and calculating and presenting aggregate results. Absolutely no score is reported in the Online Reporting System until it passes all of the QM system's validation checks.

6. TEST SETTINGS, ACCOMMODATIONS, AND SPECIAL CODES

6.1 ONLINE TESTING FEATURES AND TESTING ACCOMMODATIONS

The SAGE test delivery system provides a range of accessibility tools and accommodations for reducing construct-irrelevant barriers to accessing test content for virtually all students. The range of accommodations provided in the online testing environment far exceed the typical accommodations made available in paper-based test administrations, which were typically limited to large print, Braille, and English and foreign language audio translations. Table 4 lists the accommodations and accessibility supports currently available for the SAGE assessments.

Table 4. Accommodations and Accessibility Supports

Accessibility Feature	Description
Text-to-Speech—Directions, Passages, Items	Computer reads text and graphics aloud on directions, passages, and items. What is read and how it is read is configurable.
Text-to-Speech—Graphic Description	Computer reads graphics and tables aloud.
Magnification Interface	Student can zoom in and zoom out the entire page. This capability persists throughout the test.
Magnifier	Student can magnify a selected portion of an item.
Variable Font Size	The number of levels (generally, five levels) and rate of increase (generally, 1.25x the previous level) are configurable.
Refreshable Braille/ Tactile with External Embosser Printer	Items can be rendered to desktop embossers that can integrate Braille and tactile graphics. The items will simultaneously render on a reader-accessible screen, and the student will be able to navigate to response spaces to provide answers.
Reverse Contrast	Background turns to black, while text turns to white.
Administrator-Selectable Variable Font and Background Colors	Any foreground and background color can be supported.
Color Overlay	Any color can be laid on the screen. This persists throughout the test.
Increased White Space	This is the streamlined interface.
Sign Language—Directions, Passages, Items	This capability consists of recorded videos on sign language. Avatars are not recommended by hearing-impaired experts because they do not translate well to American Sign Language.
Translations	Versions are available in alternate languages.
Keyword Translation	This enables translators to associate keyword translations.
Glossaries and Dictionaries	These enable content developers to associate additional content with words or phrases. The content can be of multiple types, and the content shown to a student can be controlled by his or her personal profile.
Alternate Language Glossaries and Dictionaries	These enable content developers to associate alternate-language content with words or phrases. The content can comprise multiple types, and the content shown to a student can be controlled by his or her personal profile.

Accessibility Feature	Description
Administrator-Selectable Assistive Devices Integration	Our system has a standard and a streamlined interface. Most assistive devices can work with the former, and an even wider group works with the latter. If the use of the device requires relaxation of certain security features (e.g., if suppression of pop-up windows interferes with on-screen keyboards), the system can be configured to allow the test administrator to select a more permissive mode.
Line Reader	This feature will allow a student to track the line he or she is reading.
Masking	Students can mask extraneous information on the screen.
Speech-to-Text	Speech will be converted to text and then saved in the database. (Available through compatibility with third-party assistive technology.)
Auditory Calming	A tool that plays music or white noise in the background. (Available through third-party software.)
Administrator-Selectable Zoom	Default font size can be set in advance through a file upload or user interface or at the time of testing by the test administrator. Student can zoom in or zoom out at any time.
Administrator-Selectable Large Print Font	Default font size can be set in advance through a file upload or user interface or at the time of testing by the test administrator. Student can zoom in or zoom out at any time.
Administrator-Selectable Screen-Reader	The system supports an integrated screen reader that can be configured to provide a variety of support levels, each selectable by the administrator.
Additional Time	AIR's system currently does not impose a time limit on the test. It is up to the proctor to stop a student's test or stop the entire session. However, if there are unforeseen events, such as a fire alarm, that trigger need for additional testing time, AIR's system can enable a grace period extension (GPE) for a single test opportunity or for multiple test opportunities.
Segment Breaks	AIR's system has the capability of adding test segments within a test. A test segment is made up of multiple item groups and creates a logical break between segments within a test. For example, a segment break might separate a calculator from a non-calculator segment of a test.
Recorded Audio	Computer efficiently delivers recorded audio. We are able to deliver voice-audio using only about 10 Kbps of bandwidth.
Secure Print Facility	A visual accessibility feature, the secure print facility allows the secure printing of items or passages. A student requests that a passage or item be printed. The request is then encrypted and sent securely to the proctor. The proctor needs to approve the request before it is sent to the printer. In addition, this feature also allows for the delivery of real-time paper-based tests, including large print tests.
Test Pauses and Restarts	An attention accessibility feature, test pauses and restarts allows the test to be paused at any time and restarted and taken over many days. So that security is not compromised, visibility on past items is not allowed when the test has been paused longer than a specified period of time.
Writing Checklists	An attention accessibility feature generally for essay items, the writing checklist enables a student to check off writing guidelines from a checklist.
Review Test	Students can review the test before ending it.
Area Boundaries	An agility accessibility feature, area boundaries for mouse-clicking multiple-choice options allow students to click anywhere on the selected response text or button.
Language	Any language that is necessary can be supported.
Help Section	A reference feature, the Help section explains how the system and its tools work.
Performance Report	A reference feature, a performance report is available at the end of the test for the student.

6.1.1 Online Testing Features for ALL Students

In SY2016–2017, the following features were available for *all* students to access. For specific information on how to access and use these features, refer to the *Test Administration Policies and Procedures.*

- Zoom in: Students could make test questions, text, or graphics larger by clicking the Zoom icon, which has four levels of magnification.

- Highlight: Students could annotate passages or sections of passages and test questions by highlighting. Note: for SY2015–2016 the highlight tool was changed to provide the capability for students to re-apply student highlighting after pause/resume of a test.

- Pause: Students could pause the assessment and return to the last unanswered question. If an assessment was paused for more than 20 minutes, a student would not be allowed to return to previously answered test questions, except for on the ELA writing test, for which students can return to their prior response up to the point of submission.

- Cross-out response options: Students could cross out text in multiple-choice answer options by using the strikethrough function.

- Mark a question for review: Students could mark test questions they had answered to review them later (however, if an assessment is paused for more than 20 minutes, students were not allowed to return to marked test questions that were previously answered).

- Text-to-speech: Students were able to listen to test questions through a text-to-speech (TTS) feature (which required headphones). The text-to-speech feature was enabled automatically for all students in TIDE. This capability allowed students to have test instructions, writing passages, questions, or answer options read or reread to them. TTS is not available for reading passages. This procedure ensured that all students had the test questions and response options read aloud electronically using the same tone of voice and inflection for all words and symbols included in charts, tables, and graphs.

- TTS settings inside the test: The student was able to change the TTS volume, rate, and pitch while within the test on the secure browser.

- Audio scrubber for all embedded audio: Enabled students to be able to start a sound recording at any location in the audio file by moving a progress indicator in the tool. This was available for listening audio files.

- Descriptive audio: Students were able to listen to audio descriptions of interactive answer spaces in test questions. This audio was provided in addition to text-to-speech as a test setting.

- Dictionary: Students could look up definitions and synonyms in the Merriam-Webster dictionary or thesaurus. Additionally, this tool provides a Spanish translation if needed.

- Item Response Time Machine – Provides student the ability to recover a saved (or auto-saved) draft of his or her essay response.

- Mouse pointer tool – Provides student the option of mouse pointer sizes and colors.

- Calculator – Where applicable, the Desmos calculator is available to students.

- Update TDS Item Skin and Item Tools menu – In the past tools were accessible via the right-click context menu and/or the buttons at the top of the item. Younger students had a hard time with this because options were hidden from them unless they knew that the right-click was available. For the SY2015–2016 administrations, the tools were all consolidated in a single location under an industry-standard menu button at the top right corner of all items.

6.1.2 Accommodations for Special Populations

An accommodation was provided for English language learner (ELL), Individuals with Disabilities Education Act (IDEA)-eligible, or Section 504 students. An accommodation is a practice or procedure in presentation, response, setting, timing, or scheduling that, when used in testing, provides equal access to all students. State-approved accommodations do not compromise the learning expectations, constructs, grade-level standards, or measured outcome of the assessment.

To obtain detailed information about the official state policy for assessment accommodations, examine the *Accommodations Policy*.

The SAGE summative, interim, and benchmark assessments allowed for six accommodations. Each accommodation must be documented in the student's Individual Education Program (IEP), ELL, or Section 504 plan.

- Braille/Tactile Graphics: Braille/tactile graphics and embossing were available on request for students with disabilities. Text and graphics could be provided via an embosser in the student's school, which could print text and/or graphics depending on student need. More information was provided in the *Braille Requirements and Testing Manual.*

- Calculation Devices (Grade 6 Mathematics Only): Students on an IEP or 504 plan could bring in a handheld calculator to use ONLY on the calculator segment.

- Standard-Size Paper: The *Print on Request* tool provided access to schools for printing test questions. All printouts were to be securely destroyed following test sessions.

- ▪ Large-Print Paper: The *Print on Request* tool, when combined with enlarged *Print Size,* provided access to schools for printing large-print test questions. All printouts were required to be securely destroyed following test sessions.

- ▪ Scribe: A scribe was allowed for students with disabilities and for students as necessary due to temporary injury just prior to assessment as non-accommodation.

- ▪ Sign Language: The Sign Language accommodation was only available on the listening passages for ELA. Sign language could be provided via an embedded video of a human signing for students with disabilities. Human signing on any portion of the ELA, math, or science assessments was allowed ONLY for directions.

English Language Learners (ELLs)

The federal NCLB Act Title III and state U-PASS legislation and policy determine which ELL students take the SAGE summative assessments. This determination was based on the student's number of years in the United States. The pertinent guidance related to these policies is summarized below. All assessment decisions related to SAGE administrations regarding ELL students could be informed by the school's ELL team. A change was made that the ELL flag was provided in the UTREx student file. In prior years it was entered by users. USBE also created a district allowing these students to take the SAGE test if they desired. By taking the test in this special district, students' scores did not count toward their home district so they could not impact accountability.

Table 5. ELL Student Test Eligibility

Years in United States/Proficiency Level	ELA	Math	Science
Less than 1 full year Enrolled on or after April 15 of the current year	No	No	No
Less than 1 full year Enrolled before April 15 of the current year	No	Yes	Yes
1–2 years Enrolled on or after April 15 of the previous year	No	Yes	Yes
1–2 years Enrolled before April 15 of the previous year	Yes	Yes	Yes
3 years or more	Yes	Yes	Yes

Special Codes

Special codes were used, for summative tests only, to explain why a student did not participate in an assigned assessment or why a student participated in the assessment non-standard way. Non-participation codes removed eligibility so a student marked with a non-participation code would not be able to take the associated test.

When special codes were used, only one code could be selected. A description of special code usage is provided below.

Basic rules:

- Special codes were used for non-standard administrations and to explain non-participants.

- A blank assessment could be interpreted as "not participating," and a special code must be marked for the assessment.

- Less than six questions answered on an assessment was interpreted as "not participating," and a special code was required to be marked for the assessment.

- Special codes were audited for appropriate use. All student data were used for scoring, reporting, and accountability.

For special codes, select one of the following:

- "Parental Exclusion" if a parent or legal guardian has requested in writing that the student not take the test. This is the only special code that prevents a student from testing if assigned.

- "Absent – Did Not Test" if the student was not present during any part of the test administration period and was not able to make up the test.

- "Student Refuses to Test" if the student chose to give up during testing or refused to start the test.

- "Excused Medical Emergency" if the student was unable to test during the testing window due to an unanticipated medical circumstance.

- "Course Will Continue Next School Year" if the student is enrolled in a course that began this school year but will be completed in the next school year.

- "Course Instruction Not Aligned with Course Code" if the student was enrolled in a course using a course code that does not align to the instruction given.

- "Test Has Already Been Taken" if a student has already taken this SAGE test in a previous administration and therefore is not required to retest.

- "USBE Excused – Approval Needed" may only be used if USBE Assessment and Accountability staff approves in writing.

- "Student Transferred Before Testing Window" if a student was enrolled in a SAGE assessed course, but transferred out of the course before the test was scheduled to be given.

- "Accommodated" if accommodations are provided.

- "Modified" if a modified assessment was administered.

REFERENCES

Olson, J. F., & Fremer, J. (2013). TILSA test security guidebook. Washington, DC: Council of Chief State School Officers.

Wollack, J. A., & Fremer, J. (Eds.). (2013). Handbook of test security. New York: Routledge.