Computer-Based Testing Tryout Report

Contract #5417

Prepared for the California Department of Education by Educational Testing Service

April 24, 2013

Listening. Learning. Leading.®

Table of Contents

Tables

Executive Summary

This report summarizes the design, administration, and evaluation of the Standardized Testing and Reporting (STAR) Computer-based Testing (CBT) Tryout that was administered in October 2012.

The primary purpose of the STAR CBT Tryout was to obtain feedback from students, schools, and local educational agencies (LEAs) on their preparedness to administer tests on computer and to gain some information about items administered on the computer, both traditional multiple-choice (MC) items as well as new technology-enhanced (TE) item types, including multimedia items utilizing interactive animations, multiple-select response items incorporating drag-and-drop and hotspot functionalities, as well as virtual graphing items. This information will be used to inform future plans regarding the transition to CBT.

Three science tests—grade five science, grade eight science, and high school Biology—were developed for the CBT Tryout based on the Framework for K–12 Science Education and the California content standards for science. Each was administered from October 1–18, 2012, to a sample of 193 schools that were selected to include the range of demographic characteristics and technological capabilities found in California schools.

This report is organized into three main sections. The first section, *Test Design and Data Collection Methods,* describes the test design and development process as well as the test administration activities and procedures. The second section, *Results on Participation, Test Administration, and Artificial Intelligence Scoring,* summarizes the results of the pre- and post-test surveys of students and administrators as well as the associated site visits. This section includes information about the student, school, and district readiness for the CBT, as well as information collected about the school testing environment, security and processing issues, technological issues, the adequacy of training and proctoring, and the reactions and activities of test takers during the administration. Also presented are the results from the artificial intelligence scoring process.

The final section, *Psychometric Studies,* describes the results of psychometric studies evaluating the statistical properties of the items and test forms; dimensionality of the CBT forms; possible differential impacts of the CBT on student subgroups of interest at both the item and test levels; and the major factors comprising technology readiness, which may have impacted student CBT performance.

When the STAR CBT Tryout was announced, 609 LEAs expressed an interest in participating in the tryout, representing 40 percent of all LEAs in the state. This suggested a high level of statewide interest in CBT. Ultimately, 133 of the invited LEAs chose to participate, resulting in a testing sample of 193 schools. The LEAs and schools that participated covered the spectrum of preparedness for CBT ranging from those that were uncertain of their preparedness to those that indicated they were very well prepared. Though every effort was made to obtain a representative sample and to meet sample targets, students who actually took the tests represented only a small proportion of the student population in the state and generally represented only schools that had the technology infrastructure and time to participate. Therefore, caution is advised in drawing broad conclusions about CBT in the state as a whole.

Overall, the results indicate that many of the LEAs (approximately 85 percent) who participated in the tryout and responded to the survey questions believe they are ready to begin phasing in CBT for future assessments, assuming that system requirements similar to those for the CBT Tryout would apply. Additionally, the CBT Tryout demonstrated that TE items, including those with animation, can be successfully administered on the typical computer and network systems that currently exist in California classrooms.

At the same time, a significant number of LEAs in California still have either perceived or real technological barriers to CBT. Of the 358 LEAs that declined to participate, nearly half cited "we don't think the right technology is in place" as a reason for declining. Additionally, because the CBT Tryout tested only a subset of students at each participating school, no conclusions can be drawn about how prepared these schools would be to test every student on a computer, nor about the length of the testing window required to complete such testing.

Analyses of test data show that the CBT Tryout test forms provided a reliable measure of student performance and that the administered test forms, comprised of MC items with a modest number (33 percent) of TE items, resulted in a unidimensional assessment (i.e., measures a dominant construct of interest). In addition, results investigating any potential differential impact on student subgroups suggest that the CBT may lead to small differential impact at the item and test level for some subgroups. However, these findings should be interpreted with caution due to the limitations of the study, including considerations of motivation (no-stakes), timing (i.e., students were administered the CBT Tryout forms several months after they completed the course in the subject), and the design of the study, which did not include the counter-balanced administration of paper versions of the items and forms.

Finally, analyses of the student CBT readiness survey data suggest that students with more exposure to computers either in an academic or nonacademic environment and students with higher efficacy and a positive attitude toward using computers tended to perform somewhat better than expected on the assessments that comprised the CBT Tryout than students with less of these characteristics. Analyses of the school-level survey indicate that students in schools where teachers have higher levels of experience with technology also performed better on this CBT compared to their counterparts.

Introduction

Purposes of the STAR CBT Tryout

The CBT Tryout was designed to assist the State Board of Education (SBE) and the CDE in assessing California's preparedness for CBT, to identify important considerations and potential obstacles to CBT, and to better inform the future of testing in California. The CBT Tryout also provided participating students, schools, and LEAs with an opportunity to try out computerized testing, including TE item types, in a low-stakes environment. While preliminary percent-correct scores for each student were instantly delivered to schools and LEAs to demonstrate the capability of instantaneous electronic scoring, these scores were not used for any type of accountability measures.

Specifically, the planning and development of the CBT Tryout were driven by the following goals:

1. Collect feedback and data that may be useful to the CDE and the SBE in transitioning the state to CBT.
2. Provide schools and districts an opportunity to try CBT within the context of the existing STAR infrastructure with no consequences for students or school accountability.
3. Develop and administer new TE item types that can assess student knowledge in science standards not easily assessed using paper-based test items and evaluate the feasibility of artificial intelligence scoring.
4. Examine the psychometric properties of items comprising the CBT test forms, evaluate the possible differential impacts for student subgroups of interest, and determine potential factors that may have impacted student and school readiness for CBT.

Organization of the Report

The report is organized in three main sections as follows:

- *Test Design and Data Collection Methods* describes the process used to develop the test forms, the sampling and recruiting procedures, the test administration procedures, and results from the preselection survey, site visits, and a post-test survey. In addition, details on the collection and scoring of student response data, including artificial intelligence scoring methods, are also provided.
- *Results on Participation, Test Administration, and Artificial Intelligence Scoring* summarizes the LEA, school, and student participation results as well as the results of the preselection survey, site visits, and post-test survey, and the performance of the artificial intelligence scoring system.
- *Psychometric Studies* summarizes the data, methods, and results for the following analyses: the statistical properties of the CBT Tryout forms at the test and item levels; dimensionality analyses to explore the underlying structure of the CBT Tryout tests; possible differential impact of CBT administration on subgroups of interest; and major factors that may have impacted students and school readiness for CBT.

Test Design and Data Collection Methods

Test Design and Development

The CBT Tryout science tests were based on the Framework for K–12 Science Education and the California content standards for science—the same resources that guide the development of the California Standards Tests (CSTs) for science—and included a review and general alignment of core competencies to the Framework for K–12 Science Education.

ETS test developers who were already experienced in working on science assessments within the STAR Program reviewed the California science content standards and aligned those standards where they fit to the Framework for K–12 Science Education. With the CDE's approval, this information was then used to design and develop test blueprints, which also targeted California standards that historically have proven to be difficult to assess on traditional multiple-choice assessments. Specific standards were also targeted that could best be accessed through the interactive nature of a technological environment and delivery platform.

ETS and the CDE presented the draft blueprints for STAR Assessment Review Panel (ARP) review in April of 2012 and utilized the expertise and recommendations of this group of California educators to shape the development of the assessments. Using these blueprints, one CBT Tryout form was developed for each of the three CBT science tests: grade five science, grade eight science, and high school Biology.

Each test form included 60 items: 40 MC items that were a direct repurposing of the existing paper-and-pencil CST items and 20 technology-enhanced TE items that were developed to take advantage of technological capabilities such as simulations. All of the items, including the TE items, were scored dichotomously (i.e., incorrect or correct, 0 or 1, respectively), thus yielding a maximum of 60 raw score points. One hundred and sixty-six of the total number of CBT Tryout items were designed to be scored instantaneously during administration; the other 14 items were designed for post-administration artificial intelligence (AI) scoring. Instantaneously scored items included both MC items and items with finite numbers of correct responses scored using lookup tables. Items requiring AI scoring were short-answer constructed-response items traditionally requiring human scoring; these items were scored after test administration.

Similar to the regular CST science tests, each CBT form was divided into approximately two equal parts, and each part was required to be administered in a single sitting. Administration of the parts was untimed, but each part was designed to take approximately 60 minutes. Schools had the option of administering both parts in a single day, with or without a break, or on consecutive days.

Following initial item and form development, the CDE reviewed the materials to ensure they adhered to established STAR Program requirements. A second ARP meeting was convened in May 2012 to confirm alignment of the new TE items to the California science standards.

In September 2012, the CDE conducted a final review of each of the three completed online test forms and approved them for administration to students during the October CBT Tryout.

Sampling and Recruiting

Tasks leading up to the testing administration included administering a preselection survey assessing the interest of LEAs in participating, selecting a sample of participating schools, and providing training, materials, and tools to schools and LEAs. These activities are described in greater detail in this section.

The CBT Tryout was scheduled for fall 2012 and was intended to include 30,000 students: 10,000 for each of the three science test forms—grades five and eight and Biology. To ensure that participants had the opportunity to learn the content standards included in the assessment, forms were administered to students who had taken the respective science class the preceding academic year—that is, the grade five form was administered to grade six students, the grade eight form was administered to grade nine students, and the Biology form was administered to students who had completed a Biology course in the previous academic year.

Following an official announcement letter from the CDE to STAR coordinators and district superintendents, formal recruiting began on June 8, 2012, with the distribution of the preselection survey designed to assess participation interest by LEA STAR coordinators. The survey window was closed on July 6 after a sequence of reminder e-mails.

From a pool of 609 interested LEAs (representing 40 percent of all LEAs), 338 schools from 199 LEAs received an invitation to participate. Statewide distribution of invited LEAs is shown on the map in Figure A.1 on page 83 (Appendix A). These schools were selected to be representative of the state based on the following performance and demographic variables obtained from 2011 and 2012 STAR data:

- Number of students identified as proficient based on the 2012 STAR results data in English–Language Arts (ELA) (grades six and nine) and Biology[1].
- Urban/rural location
- District and school size
- Racial/ethnic diversity
- Special education status
- English-language fluency
- Socioeconomic status
- LEA perceptions of technological preparedness (from the initial recruiting survey)

The specific sampling procedure is described below:

Step 1. Schools were grouped into subgroups (i.e., strata) using urban/rural location and percent proficient status based on the 2012 STAR results data.

Step 2. The number of students needed from each stratum was determined through proportional allocation to the state. This was calculated by multiplying the target number of students by the percentage of students represented by the specific stratum among all strata.

[1] Performance results on the CSTs for Science in grades five and eight could not be used for stratified sampling because students would often have changed schools when they started grades six and nine in the fall (i.e., the transition from elementary school to middle school and the transition from middle school to high school). The options would be to use grades six and nine English–Language Arts (ELA) or mathematics CST results. Because grade nine students could take different mathematics tests, the CST for ELA (Grade 9) was used for sampling grade nine schools. For consistency, the grade six ELA CST was used for sampling grade six schools.

Step 3. A simple random sample of schools was picked from each stratum until the overall number of students from selected schools was reasonably close to the predetermined number in Step 2. In this case, the smallest sampling unit was the school.

Step 4. The extent to which the selected sample represented the state student population was evaluated in terms of the demographic characteristics that included urban/rural distribution, ethnic diversity, special education status, English learner percentage, economic status, and CST performance.

Step 5. Steps 1 through 4 above were repeated until a representative sample was obtained if the evaluation in Step 4 showed that a selected sample was not representative.

Step 6. For each school in the final sample, a replacement from among the nonselected schools was also identified to maintain a representative sample upon substitution. These schools served as replacements for any of the identified schools that later declined to participate.

LEAs were excluded if their eligible students were not co-located or if they had a special education focus. Only schools with a minimum of three eligible students were considered for participation. Efforts were made to ensure both K–6 elementary schools and 6–8 middle schools were represented in the grade six testing sample so that elementary schools had the opportunity to participate in the CBT Tryout.

ETS also attempted to recruit schools whose coordinators felt they were not prepared for CBT to ensure the testing sample included such schools in order to learn more about the potential challenges of implementing CBT in California. In many cases, these schools declined participation or participated up to a point but were unable to test students.

The complete list of participating schools and districts can be found in Appendix A, which starts on page 70.

Test Administration and Data Collection

The CBT Tryout included many of the same support structures used during a standard STAR administration: a testing window, support for the LEA, a formal security plan, and mechanisms for collecting feedback from the field. In addition, specialized surveys were created to collect feedback during and after the administration window to assess school and district readiness for CBT. As this was a new method of delivery, the test delivery engine underwent a week of technical trials in Sacramento-area schools in May 2012 to ensure the system could deliver the newly developed technology-enhanced items.

Test Administration Window

Testing commenced on October 1 and proceeded through October 12, 2012. Near the end of the testing window, a special request was received from Los Angeles Unified School District (LAUSD) for an extension of the testing window. This request was granted and the testing window was extended for an additional week, to October 18, with an additional 20 students tested.

Customer Support Procedures

The technical support resources provided for operational STAR administrations were also provided to school and district staff for the CBT Tryout. The STAR Technical Assistance Center (TAC) staff members received special training and were provided with protocols for handling changes in the types of support required for a computer-based test. Additionally, a new support structure was created so that STAR TAC representatives were immediately able to escalate technical issues related to the new computer-based test delivery system when needed. STAR TAC representatives were asked to log their CBT Tryout interactions and to record impressions, feedback, and comments.

Training Activities and Materials

Key training materials included a *Computer-based Testing Science Tryout Coordinator and Administrator Manual* (CDE, 2012b), role-specific checklists (CDE, 2012c, 2012d, 2012e), a series of how-to videos explaining installation and setup procedures (ETS, 2012a, 2012b, 2012c, 2012d, 2012e, 2012f, 2012h, 2012i, 2012j, 2012k, 2012l, 2012m, 2012n, 2012o, 2012q), and an online tutorial (ETS, 2012c). Two live Webcasts presented on September 14, 2012, provided comprehensive training for STAR coordinators, technology coordinators, and test administrators. These live Webcasts were recorded and made available for on-demand viewing (ETS, 2012g, 2012n).

Security Procedures

STAR TAC representatives monitored social media Web sites, including Instagram/Webstagram, Topsy, and Google +, looking for security breaches in the form of digital image postings and/or discussions of the items before, during, and after the testing window. Monitoring occurred from September 28 through October 26, 2012. Each social media Web site was viewed and researched three times daily during the monitoring period, and resulted in a total of 189 searches. STAR TAC found no breaches of security, discussions, or postings of any items, testing rooms, or testing computers during the monitoring window.

In addition to online monitoring, standard security procedures for administration of all STAR tests were followed, with an added emphasis on prohibiting the use of electronic devices during testing. STAR coordinators collected security affidavits (CDE, 2012f) from all personnel involved with the administration of the assessments. All administrators received training program guidelines and

materials, including the *CBT Manual* (CDE, 2012b), *Directions for Administration* (CDE, 2012a), and recorded Webcasts (ETS, 2012g, 2012n).

Site Visits

During the test administration, the CDE and ETS conducted site visits across the state that provided an opportunity to witness and record student and administrator reactions, including:

- Testing environment
- Security and processing issues
- Attitudes of administrators and test takers
- Technology issues
- Adequacy of training and proctoring

Post-test Survey

Two post-test surveys, one for LEAs and schools and one for students, were created in consultation with experts in K–12 CBT administration requirements as well as survey design.

LEA and School-level Survey

Following testing, an online survey was distributed to districts and independently testing charter school STAR coordinators, LEA technology coordinators, and school-level administrators. The survey included questions about the CBT Tryout testing experience and the technology environment at the school and LEA levels. A series of targeted e-mail reminders and follow-up phone calls were conducted to maximize survey response rates. The surveys are included in Appendix B starting on page 88.

Student-level Survey

Immediately after completing the second section of the test, students were asked to complete a survey consisting of questions about their technology readiness for CBT and the overall presentation of the test. The student post-test survey also included specific questions designed to learn more about student interactions with the new item types by means of a modified virtual cognitive lab where students responded to questions about the simulation using a 4-point Likert scale. Please note that while questions similar to those that would be asked in a cognitive lab were included, conducting a true cognitive lab—where a researcher would interact directly with the student either in person or virtually by telephone or video to elicit verbal reporting—was beyond the design, development, and logistical constraints of this CBT Tryout. This survey is included in Appendix B starting on page 84.

Scoring Student Response Data

Selected Response Item Scoring

All of the 40 multiple-choice items, as well as some of the TE items, for all tests were dichotomously scored by computer using scoring keys stored on a centralized server. Preliminary percent-correct scores on these items for all students were delivered to schools and districts instantly in electronic form.

Artificial Intelligence Scoring

The STAR CBT Tryout included a total of 14 constructed-response items that could be scored using one of two automated scoring engines: *m-rater* and *c-rater*. For the three tests administered during the CBT Tryout, seven constructed-response mathematics-related items were scored with m-rater and seven constructed-response items were scored with c-rater. A short description of each

automated scoring engine and the methods used to evaluate its performance are presented in the following sections.

M-rater

M-rater is designed to score graphs, equations, and numeric-entry responses. Items suitable for m-rater scoring are mathematical in nature; responses are determinate, meaning that the items have clearly identified and specified correct answers that are not subject to interpretation. To accomplish such scoring, m-rater uses a computer algebra system that is based on an answer key that is entered into the scoring system before administration. When the response to a task is an equation, a numeric answer, or a graph, m-rater scores the response as right or wrong with nearly 100 percent accuracy. Although m-rater can also identify specific characteristics of a response and assign a partial-credit score based on the presence or absence of these predetermined features, in the CBT Tryout, partial credit was not permitted and all constructed-response items were scored as right or wrong.

For the CBT Tryout, the m-rater scoring engine was used to score two item types: numeric entry and graph responses. The m-rater scoring engine can accept a range of answers for numeric entry items and can recognize different ways to express the same answer. Similarly, the graph key can be used to score student-generated plots and take into account specific tolerances for plotting accuracy. In the CBT Tryout, such tolerances were employed to give students the benefit of the doubt when plotting answers, allowing for the different eye-hand coordination capabilities of students when using a mouse. That is, students were given credit for the correct answer when they were able to click within close proximity of the correct response on a grid, as opposed to pointing at the exact location of the correct answer.

M-rater Evaluation Methods

Since m-rater is used to score tasks with clearly determined answers, human-to-machine scoring agreement is not necessary. Instead, potential errors in m-rater results are detected by processing the student response data and calculating all possible answers using ancillary software tools outside of m-rater. Previous research (Leacock & Chodorow, 2003) has shown that when there are differences between humans and m-rater scoring, the root cause of the disagreement can most often be attributed to: 1) human error; 2) complicated rubrics that are difficult for humans to interpret but can be programmed into computers; 3) humans forgiving typographical errors not accepted by m-rater; and 4) complications with the computer interface due to either supplemental entries being permitted or incorrect deciphering of results for m-rater.

C-rater

C-rater is designed to score short text responses for correct content and uses natural language processing to assess if specific concepts are present in the response. C-rater assesses responses for content by first creating model sentences with required words. Then, c-rater creates synonyms for alternate ways to write the same response. In scoring student responses, each response is compared to the model sentences and the alternate responses using the synonyms. Once a student's response is associated with a model sentence with synonyms, scoring rules are used to assign points to the specific concepts that are present in the response.

C-rater Evaluation Methods

C-rater scoring models can be evaluated by two types of measures: measures of association that are descriptive in nature and measures of agreement that indicate if the models are overscoring or underscoring student answers compared to humans. The measures of c-rater models associated with human-generated scores are typically correlations and quadratic-weighted kappa statistics; both measures use a guideline threshold of 0.70. This value was selected on the conceptual basis that it represents about half of the association between two raters: human and human, or human and machine. Given that c-rater is compared to human scoring, the expected performance of c-rater is bound by the

consistency of human scoring; if the interrater agreement between independent human raters is low (e.g., less than the 0.70 threshold), then automated scoring is disadvantaged in demonstrating an acceptable level of performance. This is because of the potential lack of reliability in the human scoring process upon which c-rater is modeled and evaluated, rather than any particular failing of automated scoring.

When assessing the agreement of c-rater scoring, c-rater performance is compared to human scoring. Two measures are used to determine any reduction in agreement: degradation with a guideline threshold of ≤ 0.10, and the effect size of any difference in mean scores (standardized mean score difference with a guideline threshold of ≤ 0.15). Degradation is calculated as the difference between the correlation or kappa between double-human scoring minus the correlation or kappa between human scoring and c-rater scoring. The mean score differences indicate any disparities between human scoring and c-rater scoring in either awarding or denying credit for correct answers. Additional evaluations such as association with external variables, other test part scores, and subgroup difference analyses (e.g., gender, ethnicity) are also completed depending on the availability of data.

All the performance criteria are applied to the independent evaluation sample used to validate the scoring models. The results from the evaluation sample, which is separate from the model-building sample, represent a more generalizable measure of performance expected to be consistent with what would be observed from future data.

Results on Participation, Test Administration, and Artificial Intelligence Scoring

Results on Participation and Test Administration

This section of the report presents a summary of:

1. LEA, school, and student participation and demographic representation;
2. Missing and incomplete data;
3. Pre- and post-test survey results;
4. Technical support and system performance
5. Customer support; and
6. Technical analyses of the test delivery system.

LEA, School, and Student Participation Summary

The targeted number of test takers for this study was 30,000 students, although only 21,473 students took part in the CBT Tryout. The participation rates are summarized in Table 1. As described in the *Sampling and Recruiting* section on page 9, invited schools that declined participation were replaced with substitute schools to help maintain a representative sample. Of all the LEAs and schools that received an invitation to participate, 175 LEAs and 285 schools either agreed or had not responded by the time the deadline was reached for replacing schools. In total, 54 schools that declined were replaced with substitute schools that had matching demographic characteristics.

A total of 221 schools from 136 LEAs registered students by uploading Pre-ID files (see row three in Table 1) and a total of 21,473 students from 193 schools and 133 LEAs participated in the CBT Tryout. Although not specifically targeted in the sampling plan, the participating schools included 12 charter schools, of which four were independently testing schools. No SBE charters participated in the CBT Tryout.

Table 1. Summary of Participation in the STAR CBT Tryout

	No. LEAs	No. Schools	No. Students
Were invited to participate, including schools that canceled and were replaced	199	338	69,753
Planned to participate or had not yet responded by the time recruiting closed and training began in September	175	285	58,481
Used the CBT Tryout system and uploaded Pre-ID files	136	221	31,546
Tested students	133	193 *	21,473

* Of the 193 testing schools, 12 were charter schools. Of the 12 charter schools, 4 are independently testing schools under the STAR Program.

A number of different factors most likely accounted for the attrition of schools that uploaded Pre-ID files but did not participate. Although schools were selected from a pool of LEAs that indicated a willingness to participate, the actual schools within these LEAs may or may not have expressed the same interest or ability to participate. Additionally, many schools and LEAs cited factors such as scheduling conflicts with other testing windows, vacation schedules, or the unavailability of computer labs as reasons they were not able to participate. Finally, some schools and LEAs stated they needed more advance notice to be able to participate. For example, the need for advance notice was essential to LAUSD, which was unable to test a large number of students as a result of scheduling conflicts that could not be resolved between the notification and administration of the CBT Tryout. Although the CBT Tryout was constrained to a two-week testing window, any future CBT testing would benefit

from a longer testing window, potentially increasing participation rates as a result of fewer scheduling conflicts and allowing more test takers to cycle through a limited number of computers.

Analysis of Statewide Demographic Representation

The total number of LEAs, schools, and students for each of the three CBT Tryout tests with valid response data is summarized in Table 2. As noted in the table, grade six had the highest participation rate, with over 96 percent of the targeted 10,000 students participating. Conversely, Biology had the lowest participation rate with only 54 percent of the targeted 10,000 students participating.

Table 2. Summary of LEAs, Schools, and Students

	No. *			% of 10,000 Target
Science Grade	LEAs	Schools	Students	
Grade 6	84	108	9,659	96.1%
Grade 9	40	43	6,032	60.3%
Biology	53	56	5,404	54.0%

* Note: Participating LEAs and schools may have participated in more than one grade-level test; therefore, some are counted more than once in this summary.

The breakdown of participating students by geographical location and test is summarized in Table 3 and Figure 1 below, which graphically depicts the range of participants by county. Northern and southern California are defined by the horizontal line drawn in Figure 1 along the northern borders of San Luis Obispo, Kern, and San Bernardino counties. As noted in both Table 3 and Figure 1, the sample included students from across the state, with the largest amount of participation from southern California.

Table 3. Geographic Breakdown

Science Grade	Region	No. LEAs	No. Schools	No. Students
Grade 6	Northern California	36	47	3,836
	Southern California	48	61	5,823
Grade 9	Northern California	16	17	1,853
	Southern California	24	26	4,179
Biology	Northern California	29	30	2,969
	Southern California	24	26	2,435

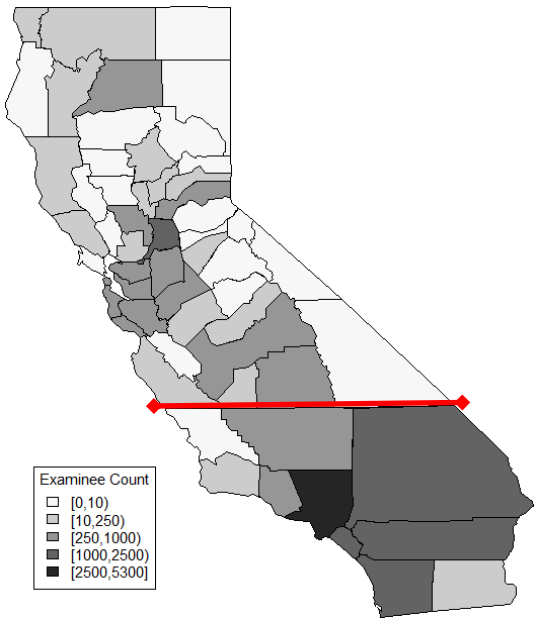


Figure 1. Examinee Participation Counts by County

The demographic information and CST performance of the participants compared to the total student population for each grade is listed in Table 4 on the next page. For comparison purposes, demographic characteristics for the participants in the CBT Tryout are based on the results data collected for the 2012 CSTs for grades five and eight and Biology, respectively. As a reminder, each of the participants had taken the related CST content earlier in the spring in their previous grade/course. For example, the grade six participants in the CBT Tryout in the fall had completed the grade five science CSTs earlier in the spring of 2012.

As illustrated in the table, the participants were reasonably representative of the population, including all of the subgroups of interest described earlier in the *Sampling and Recruiting* section. In most cases of difference between the subgroup sample and the population, the percentage of students in smaller subgroups was greater than the percentage within the overall population. For example, 3 percent of the grade six participants were identified as "small town/rural" compared to 2.4 percent in the population.

Of note, 11 percent (n=679) of the grade nine students could not be matched to a Statewide Student Identifier using the 2012 grade eight testing data—these students are listed as "unknown" in Table 4. One possible explanation for this may be related to the shift in student population from one school to another, which occurs in the middle-to-high-school transition.

Table 4. Student Demographic Characteristics for the CBT Tryout

	Grade 6				Grade 9				Biology			
	Population		Sample		Population		Sample		Population		Sample	
CST No.	428,868		9,659		434,601		6,032		555,736		5,404	
CST mean	367.9		370.5		395.7		407.0		354.9		360.2	
CST SD	66.2		65.7		98.8		97.6		61.1		54.2	
	No.	**Pct**	**No.**	**Pct**	**No.**	**Pct**	**No.**	**Pct**	**No.**	**Pct**	**No.**	**Pct**
CST percent proficient *	255,888	59.8	9,014	61.6	288,731	66.7	5,137	71.3	288,813	52.0	4,911	57.3
Male	214,341	50.0	4,688	48.5	217,858	50.1	2,709	44.9	274,777	49.4	2,597	48.1
Female	214,414	50.0	4,689	48.5	216,623	49.8	2,643	43.8	280,784	50.5	2,681	49.6
Gender unknown	113	0.0	282	2.9	120	0.0	680	11.3	175	0.0	126	2.3
American Indian	2,571	0.6	57	0.6	2,936	0.7	54	0.9	3,769	0.7	64	1.2
Asian American	38,905	9.1	1,183	12.2	38,132	8.8	630	10.4	53,653	9.7	372	6.9
Pacific Islander	2,387	0.6	47	0.5	2,620	0.6	23	0.4	3,145	0.6	38	0.7
Filipino	11,770	2.7	290	3.0	12,922	3.0	140	2.3	17,397	3.1	133	2.5
Hispanic	223,580	52.1	4,892	50.6	221,653	51.0	2,675	44.3	277,551	49.9	2,684	49.7
African American	27,365	6.4	483	5.0	29,087	6.7	281	4.7	36,979	6.7	301	5.6
White	110,977	25.9	2,238	23.2	117,481	27.0	1,441	23.9	150,452	27.1	1,584	29.3
Two or more races	11,313	2.6	187	1.9	9,770	2.2	109	1.8	12,790	2.3	103	1.9
Ethnicity Unknown	0	0.0	282	2.9	0	0.0	679	11.3	0	0.0	125	2.3
No special education services	406,223	94.7	8,595	89.0	414,946	95.5	4,931	81.7	516,112	92.9	5,031	93.1
Special education services	22,645	5.3	782	8.1	19,655	4.5	422	7.0	39,624	7.1	246	4.6
Special education unknown	0	0.0	282	2.9	0	0.0	679	11.3	0	0.0	127	2.4
English Only	238,457	55.6	4,969	51.4	243,707	56.1	2,871	47.6	307,879	55.4	3,010	55.7
Initially fluent English proficient	19,705	4.6	475	4.9	36,703	8.4	471	7.8	44,334	8.0	354	6.6
English learner	93,779	21.9	2,352	24.4	55,462	12.8	773	12.8	70,003	12.6	554	10.3
Reclassified fluent English proficient	76,433	17.8	1,575	16.3	98,055	22.6	1,235	20.5	132,695	23.9	1,356	25.1
English proficiency unknown	494	0.1	288	3.0	674	0.2	682	11.3	825	0.1	130	2.4
Not economically disadvantaged	165,592	38.6	3,433	35.5	180,801	41.6	2,015	33.4	256,082	46.1	2,214	41.0
Economically disadvantaged	262,687	61.3	5,935	61.4	253,099	58.2	3,335	55.3	298,917	53.8	3,061	56.6
Unknown economic status	589	0.1	291	3.0	701	0.2	682	11.3	737	0.1	129	2.4
Metropolitan	418,352	97.5	9,369	97.0	423,657	97.5	5,697	94.4	544,463	98.0	4,962	91.8
Small town/Rural	10,339	2.4	290	3.0	10,749	2.5	335	5.6	10,933	2.0	442	8.2
Urban/Rural unknown	177	0.0	0	0.0	195	0.0	0	0.0	340	0.1	0	0.0

* CST percent proficient for grade six was determined using results for the 2012 CST for Science (Grade 5). CST percent proficient for grade nine was determined using results for the 2012 CST for Science (Grade 8). CST percent proficient for Biology was determined using the 2012 CST for Biology results.

Discussion of Missing and Incomplete Data

There were a number of incomplete records in the data set due to students not ending the test properly or not taking one or both parts of the test for reasons such as school technology failure or student absence. These records were flagged as invalid start and/or end times by part. For each part, an invalid start time flag was assigned if the testing session failed to start, and an invalid end time was assigned if the testing session did not end properly. Students identified with both valid start times and invalid end times include those students who were unable to complete at least part of the test form and/or the post-test survey as expected and those who were able to respond to all test questions but might not have followed the proper procedure to end the testing session.

The summary of the number of invalid start and end times along with a summary of items completed by test part is included in Table 5. As illustrated, there were no students with both an invalid start and end time for Part 1 in any of the three tests. In contrast, there were relatively larger numbers of students with both invalid start and end times in Part 2. These counts include students who were absent from testing as well as those who were not able to log onto the second testing session. Both parts included cases where students had a valid start time but an invalid end time; however, there were notable differences. As noted in Table 5, most students who had a valid start time but an invalid end time for Part 1 completed one or more items; however, most students with a valid start time and invalid end time for Part 2 left all items in that part blank (i.e., missing). There were no cases with an invalid start time and valid end time.

Items in all-blank cases or items that appeared after the test unexpectedly ended were treated as "not presented" for the purpose of analysis in *Psychometric Studies,* starting on page 35.

Table 5. Number of Cases with Invalid Start or End Time by Part for Each Test

		Invalid Start and Invalid End (All Blank)	Valid Start and Invalid End		
			Completed All Items	Unable to complete One or More Items	All Blank
Grade 5	Part 1	0	12	42	3
	Part 2	364	6	66	211
Grade 8	Part 1	0	1	20	1
	Part 2	212	22	65	107
Biology	Part 1	0	2	6	15
	Part 2	147	24	94	112

Note: For the grade five science and grade eight science tests, Part 1 included items 1–30; Part 2 included items 31–60. For the Biology test, Part 1 included items 1–31 and Part 2 included items 32–60.

Preselection (Initial) Survey Results

A total of 968 LEAs responded to the initial preselection survey, accounting for approximately 60 percent of all California LEAs. The results that follow are a summary of these responses. Note that several of the survey questions asked respondents for their personal perceptions on how prepared their LEAs were for the CBT Tryouts. These survey respondents were provided with information on the minimum system requirements (ETS, 2012p) but no other specific criteria or tools were provided for evaluating preparedness of the CBT.

A majority (63 percent) of LEAs that responded to the preselection survey indicated that they wanted to be considered for participation in the STAR CBT Tryout (see Figure 2).

Figure 2. Would you like to be considered for participation in the CBT Tryout?

Of the responding LEAs, 53 percent indicated that they had good or excellent technical support (see Figure 3).

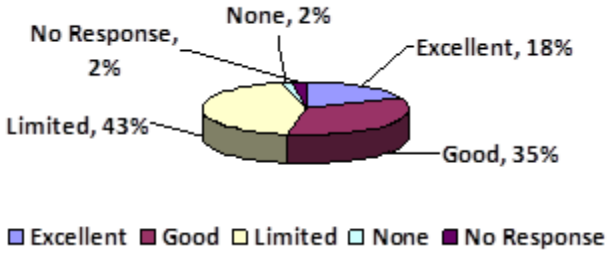


Figure 3. Please rate the availability of technical support for computing at schools in your district.

According to the personal opinion of the respondents, a large majority, 70 percent, indicated that they thought that the school district was prepared for computer-based testing (see Figure 4).

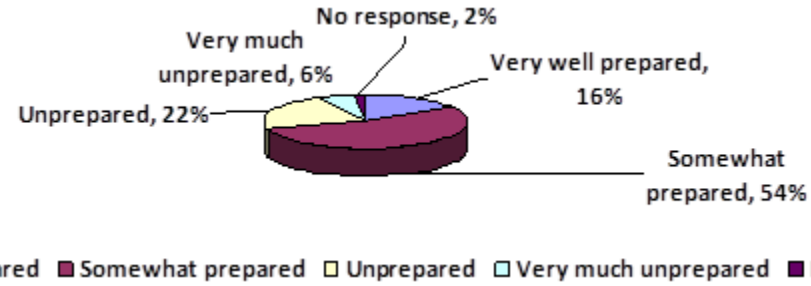


Figure 4. In your personal opinion, how prepared is your school district for computer-based testing?

Although 28 percent of the LEA respondents indicated that they thought that their school district was "Unprepared" or "Very much unprepared," these LEAs represented relatively few numbers of schools and students in the sample (see Table 6). For example, while 22 percent of LEAs responding indicated that their schools were "unprepared," that represented only 10.2 percent, 7 percent, and 10.7 percent of the grades six, nine, and Biology schools, respectively. Conversely, the vast majority of schools participating were represented by the LEAs that indicated that their schools were either "Somewhat" or "Very well prepared."

Table 6. Perceptions of Technological Preparedness

	Grade 6				Grade 9				Biology			
	Schools		Students		Schools		Students		Schools		Students	
	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct
Very much unprepared	2	1.9	69	0.7	0	0.0	0	0.0	2	3.6	190	3.5
Unprepared	11	10.2	523	5.4	3	7.0	687	11.4	6	10.7	509	9.4
Somewhat prepared	66	61.1	6,668	69.0	28	65.1	3,295	54.6	36	64.3	3,027	56.0
Very well prepared	25	23.1	2,124	22.0	11	25.6	1,943	32.2	12	21.4	1,678	31.1
Unknown	4	3.7	275	2.8	1	2.3	107	1.8	0	0.0	0	0.0

Interestingly, at the time of the survey, 84 percent of the LEAs were currently using computers for assessment purposes (see Figure 5).

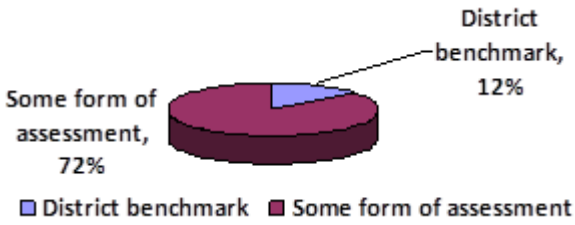


Figure 5. Do you currently use computers for assessment? (Please specify)

For LEAs that declined to participate (n=358), the most frequently cited responses provided are listed in Table 7. In this case, the respondents were asked to select as many reasons as applicable. The most often-identified reasons were related to having no staff time available (46 percent) and/or not having the right technology in place (also 46 percent).

Table 7. Reasons LEAs Declined Participating

46%	No staff time available to support a tryout
46%	Don't think the right technology is in place
21%	Need to know more before committing
17%	Don't have the authority to agree on the district's behalf
12%	District schedule conflicts with October 1–12 * testing window
19%	Other

Post-Test Survey Results

LEA and School Level Survey

Following the closure of the testing window, an online post-test survey was distributed to all participating LEA STAR coordinators, asking them to forward the survey to district-level technology coordinators and those staff members involved in the administration process within the school. Submissions were accepted from October 18 through 31, 2012, with a total of 243 completed surveys received. Of all those received, only 133 responses could be associated with a specific grade/content test. These results are summarized in the table below. A full summary of all 243 submitted surveys is available in Appendix C on page 95.

Respondents were asked to categorize their role in administering the CBT Tryout into one of three categories as defined in the training materials. The majority of surveys received were from LEA STAR coordinators and school level test administrators (see Table 8).

Table 8. Percentages of the Roles of Survey Responders

Test Administrator Role	Grade 6 (N=65)	Grade 9 (N=12)	Biology (N=56)
STAR Coordinator for your LEA	40%	17%	32%
Technology Coordinator for your LEA	11%	0%	21%
Test Administrator at the school level	28%	42%	27%
Other	5%	8%	9%
Multiple roles	14%	25%	7%
No response	3%	8%	4%

Based on the responses, the total number of students at the schools, and the total number of computers available for educational purposes is listed in Table 9. The ratio of students-to-computer was highest in grade nine, with one computer for every 11 students; and lowest in grade six, where there was one computer for every 5 students.

Table 9. Total Numbers of Students and Computers Available at Surveyed Schools

	Grade 6	Grade 9	Biology
N Students	16,083	9,461	23,427
N Computers available	3,134	834	2,578
N Ratio of students to one computer	5	11	9

The majority of respondents indicated that the CBT system was easy or very easy to use (see Table 10), Very few respondents found the system difficult to use.

Table 10. Ease of Use of the STAR CBT Tryout System

	Grade 6	Grade 9	Biology
Very Easy	17%	25%	14%
Easy	60%	33%	57%
Neutral	17%	42%	23%
Difficult	5%	0%	5%
Very Difficult	0%	0%	0%
No response	2%	0%	0%

Seventy-five percent or more of respondents indicated that the technology worked as they expected it to work (see Table 11).

Table 11. Technology Performance

	Grade 6	Grade 9	Biology
Yes	75%	92%	75%
No	23%	8%	21%
No response	2%	0%	4%

Of those reporting that the technology did not work as expected, 87 percent, 100 percent, and 33 percent in grade six, grade nine, and Biology classes, respectively, cited problems with school equipment, technology, or software as the main reason.

Based on the results of the survey, the majority of respondents found the resources provided to the LEAs helpful (see Table 12, Table 13, and Table 14 for grade six, grade nine, and Biology results, respectively).

Table 12. Resource Use—Grade Six

Grade 6	Very helpful	Helpful	Somewhat helpful	Not very helpful	Not at all helpful	Did not use/No response
Webcast	20%	28%	20%	8%	2%	23%
Manual	18%	35%	28%	0%	0%	18%
Tutorial	35%	35%	11%	2%	0%	17%
Management Tools	28%	45%	15%	3%	0%	9%
STAR TAC	28%	20%	2%	0%	2%	49%

Table 13. Resource Use—Grade Nine

Grade 9	Very helpful	Helpful	Somewhat helpful	Not at all helpful	Not very helpful	Did not use/No response
Webcast	8%	25%	17%	8%	8%	33%
Manual	25%	50%	17%	0%	0%	8%
Tutorial	33%	42%	17%	0%	0%	8%
Management Tools	50%	42%	8%	0%	0%	0%
STAR TAC	8%	25%	8%	0%	0%	58%

Table 14. Resource Use—Biology

Biology	Very helpful	Helpful	Somewhat helpful	Not very helpful	Not at all helpful	Did not use/No response
Webcast	18%	29%	14%	9%	5%	25%
Manual	23%	32%	23%	2%	0%	20%
Tutorial	32%	32%	14%	2%	0%	20%
Management Tools	25%	41%	4%	7%	0%	23%
STAR TAC	21%	18%	11%	4%	0%	46%

Nearly all of the LEA STAR coordinators reported that their schools were able to successfully administer the tests (see Table 15). The 3 percent of respondents whose response was "No" for grade six indicated that the reasons they were not able to successfully administer the test was because they "Didn't have the right technology in place."

Table 15. Successful Administration of Schools

	Grade 6 (N=32)	Grade 9 (N=4)	Biology (N=22)
Yes	94%	100%	95%
Some schools yes, some no	3%	0%	5%
No	3%	0%	0%

A significant percentage of LEA coordinators in each grade level/content area believed that they were about as prepared as they initially thought or that they were better prepared than they initially thought (see Table 16). Note, other than information about minimum system requirements (ETS, 2012p), no other specific criteria or tools were provided for evaluating preparedness for the CBT.

Table 16. Changes in Perception to LEA's Preparedness

	Grade 6	Grade 9	Biology
We are better prepared than I thought	38%	50%	32%
We are about as prepared as I thought	47%	25%	50%
We are less prepared than I thought	16%	25%	18%

As noted in Table 17, all LEA STAR coordinators indicated a preference to administer future STAR tests either on computer or on a mix of computer and paper delivery. The percentage of LEA coordinators who preferred only computer administration was higher for grade six (84 percent) than for grade nine (50 percent) or Biology (59 percent).

Table 17. Preferred Means of Delivery

	Grade 6	Grade 9	Biology
Computer only	84%	50%	59%
Paper only	0%	0%	0%
A mix of computer and paper	16%	50%	41%

For those LEA STAR coordinators who responded "computer only" or "a mix of computer and paper" as preferred means of delivery, the approximate percentage of students they would prefer to test during the first operational CBT administration is listed in Table 18.

Table 18. Roughly How Many Students Might You Prefer to Test on Computer in Your First CBT Administration?

	Grade 6	Grade 9	Biology
Ten percent of students	3%	0%	0%
Twenty-five percent of students	23%	50%	14%
Fifty percent of students	26%	25%	48%
Seventy-five percent of students	48%	25%	38%

School-level test administrators were also asked to respond to a question asking how students reacted to the CBT Tryout. Unfortunately, the majority of them did not respond (see Table 19). Of those that did respond, most reported positive reactions by the students.

Table 19. Test Administrator Perceptions of Student Reactions

	Grade 6	Grade 9	Biology
Positive	28%	50%	21%
Indifferent	3%	0%	11%
Negative	2%	0%	0%
Other	6%	8%	2%
No response	62%	42%	66%

From the response listed as "Other," the following reactions were listed:

- Loved it! They were so attentive and engaged.
- Frustrated because the program was not working appropriately.
- Positive until computers froze.
- They said they would prefer to do an online test like this instead of the paper pencil they did last year.
- There was a mixture of reaction, from positive to negative. . . it appeared easier for students to flip through the computerized assessment.
- 90% positive

Test administrators were also asked to rate the readiness of their school for CBT by considering infrastructure, computer assessment software, and administrators', teachers', and students' technology readiness. However, no specific criteria or tools other than minimum system requirements (ETS, 2012p) were provided for evaluating preparedness of the CBT. The majority of respondents indicated that they believed that their school was "somewhat" or "very well" prepared for CBT (see Table 20).

Table 20. Perception of School Readiness

	Grade 6	Grade 9	Biology
Very well prepared	3%	17%	5%
Somewhat prepared	22%	17%	13%
Unprepared	8%	25%	9%
Very much unprepared	5%	0%	5%
No response	63%	42%	68%

Table 21 shows that only a small percentage of test administrators indicated that they had prior CBT experience. Large percentage of test administrators, however, did not provide responses to this question.

Table 21. Previous CBT Experience

	Grade 6	Grade 9	Biology
Yes	15%	17%	7%
No	20%	42%	27%
No response	65%	42%	66%

For those who responded "Yes" in Table 21 above, the previous types of CBT experiences reported are listed in Table 22.

Table 22. Previous Types of CBT Experience

	Grade 6	Grade 9	Biology
District benchmarks only	10%	0%	0%
Both district benchmarks and teacher classroom tests	40%	50%	25%
Teacher classroom tests only	40%	50%	75%
No response	10%	0%	0%

Student-level Survey

A post-administration survey was also given to students. The purpose of this survey was two-fold: 1) to investigate student familiarity and comfort with electronic devices, and 2) to solicit their perceptions of the overall test presentation, item functionality, and task orientation and clarity. The survey questions are listed in Appendix B starting on page 84.

Students responded positively to questions about the test design and the computer platform. As shown in Table 23, 89 percent of the grade six students, 86 percent of the grade nine students, and 76 percent of the Biology students indicated that the instructions for completing the simulation tasks were clear and easy to understand. The students also reported that the items that used specific functionalities like drop-down lists, drag and drop, virtual timers, and running animations were easy to understand.

Table 23. Aggregate Responses by Survey Topic per Test

	Grade 6		Grade 9		Biology	
Survey Topic	**Agree**	**Disagree**	**Agree**	**Disagree**	**Agree**	**Disagree**
Instructions were easy to understand (survey questions 16 + 20)	89%	11%	86%	14%	76%	24%
Pictures were easy to understand (17 + 21)	91%	9%	91%	9%	88%	12%
Functionality was easy to use (18 + 22)	87%	13%	87%	13%	84%	16%
Task-specific response were easy to understand (19 + 23)	81%	19%	85%	15%	87%	13%

While a majority of students indicated that the instructions, tasks, and functionality were easy to understand, the increased cognitive demand of these items seemed to pose a challenge for some students. This finding was supported by their subjective opinion about of specific capstone items—the final item in a series that concludes and summarizes the simulation. For example, the survey results indicated that students in grade six and Biology found graphing more challenging than some of the other tasks (survey question 19). This was supported by a *p*-value for the grade five graphing item (item #30) of 0.06 and omit rate at 2.38 percent; and a *p*-value for the Biology graphing item (item #38) of 0.17 and omit rate of 2.41 percent. However, the overall understanding of images and functionalities were considered highly intuitive as supported by the 84 to 91 percent agreement responses in the survey, across the three grades.

Analyses of Adequacy of Technical Support and System Performance

The CBT Tryout was administered through a comprehensive online test delivery and management system that was custom configured to meet the following goals:

- Administer new technology-enhanced item types smoothly and seamlessly.
- Keep minimum system requirements as low and inclusive as possible.
- Maintain the same high level of system security expected for high-stakes STAR operational testing.
- Make system installation and operation as easy and user-friendly as possible.

Overall, information obtained from the CBT Tryout results suggests that these goals were met. However, various areas for improvement were also identified. This section summarizes technical issues that were encountered and provides suggestions for system improvement.

The following list contains a summary of all reported and confirmed technical issues:

- The autotransition feature that occurs between Part 1 and Part 2 of the test resulted in students inadvertently being advanced to Part 2 before the administrator gave permission to do so. For future testing, this could be resolved by adjusting the system so that it requires students to stop after Part 1.
- Under certain circumstances, the "pause" feature froze and required a hard restart of the testing computer. However, students were able to resume their testing beginning with the last item that they had worked on because the program was designed to save and return to the most current state and response when computer freezes of unknown origin occur. No student data were lost as a result of this freezing issue. The issue was resolved and would not be expected to reoccur for any future testing.
- At the beginning of the testing window, an incorrect technical support phone number was distributed; this was corrected as soon as it was discovered.
- In a few isolated instances, some users launched the nonconfigured contingency plan application accidentally. As a result, a small percentage—less than two percent—of test takers were exposed to tests that included items that were not displayed appropriately. Response data collected for these users were excluded from the analyses. For future testing, this could be avoided through adjustments to the program file configurations.
- Due to intermittent wireless connectivity issues, certain students could not see all of the drag-and-drop labels that were included in the last item of five that comprise the grade five simulation, which included multiple items. In addition, due to the item's unique design, the "Next" button did not appear until all the labels were placed into the drop regions. As a result, design of the item navigation features need to be reconsidered for future improvement.

- An interruption occurred during the live recording of the Webcast for STAR coordinators and test administrators several times due to technical problems; significant editing was required prior to posting the revised version. ETS is now working with a new Webcasting provider.

This list contains a summary of suggestions for CBT testing system improvement:

- Make the function for grouping students within a school by testing time and location more intuitive and user friendly, which includes making both available fields alpha-numeric and renaming the "Session" field to avoid confusion with the Session ID.
- Collect the student date-of-birth information in Pre-ID as three separate fields for day, month, and year rather than a single field.
- Make the "name" fields longer in order to accept longer names like "Christopher." (These fields were kept intentionally short to reflect the current STAR Pre-ID layout requirements.)
- Revise the *Directions for Administration* to correct some gaps in the explanation of procedures.

Tutorial Usage

The tutorial was downloaded 6,276 times. However, because the tutorial was installed and run locally—it could have been run one or multiple times after it was downloaded—the total number of users running the tutorial is unknown.

Operating System (OS) Usage Statistics, Students

The number of installations for the Test Delivery Engine on computers used by students collected between September 17 and October 19, 2012, are listed by OS version in Table 24.

Table 24. OS Usage, Students

OS Version	No. Machines	Total
Windows 2003	3	2,490
Windows 7	1,122	
Windows XP	1,298	
Windows Vista	62	
Windows Server	4	
Windows NT	1	
Mac OS X 10.8.2	12	566
Mac OS X 10.8.1	1	
Mac OS X 10.8	5	
Mac OS X 10.7.5	26	
Mac OS X 10.7.4	38	
Mac OS X 10.7.3	10	
Mac OS X 10.7.2	1	
Mac OS X 10.6.8	373	
Mac OS X 10.6.7	4	
Mac OS X 10.6.6	7	
Mac OS X 10.6.4	25	
Mac OS X 10.5.8	60	
Mac OS X 10.5.6	1	
Mac OS X 10.5.2	3	

Statistics on Browser Usage

The number of visits to the STAR CBT Web site by browser type is listed in Table 25. This information shows that while there were a number of different browsers used by technology coordinators to access the Management Tools, the most commonly used browsers were Internet Explorer and Firefox.

Table 25. Browser Usage per Visit, Technology Coordinators

OS Version	No. Visits
Internet Explorer	3,593
Firefox	1,704
Chrome	833
Safari	783
IE with Chrome Frame	38
Opera	10
Android Brower	4
Safari (in-app)	1

Statistics on OS Usage, Technology Coordinators

The total number of times that technology coordinators accessed Management Tools through the use of a specific operating system is illustrated in Figure 6. Of these, the majority of visitors were using either Windows or Macintosh operating systems.

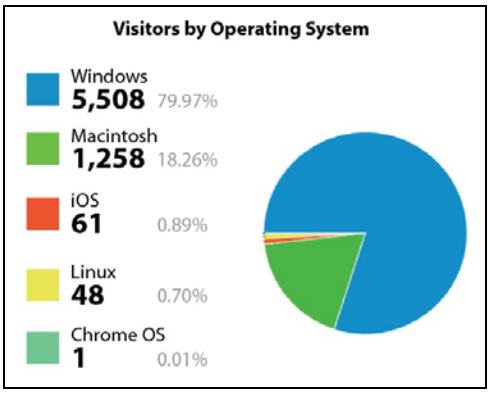


Figure 6. Statistics on OS Usage per Visit, Management Tools

Online Management Tools Visitors Overview

Information regarding visits to the star.caltesting.org Web site is illustrated in Figure 7 and indicates that there were a total of 6,888 visits and 80,913 page views, with the average visit lasting about 11 minutes.

Figure 7. Overview of Management Tools Visitors

Of these visitors, 6,279 were in California; the remaining 609 visits were out-of-state, staff-related visits. Bubbles in Figure 8 show the origins of these visits; the size and color of each bubble represents the number of visits from each location.

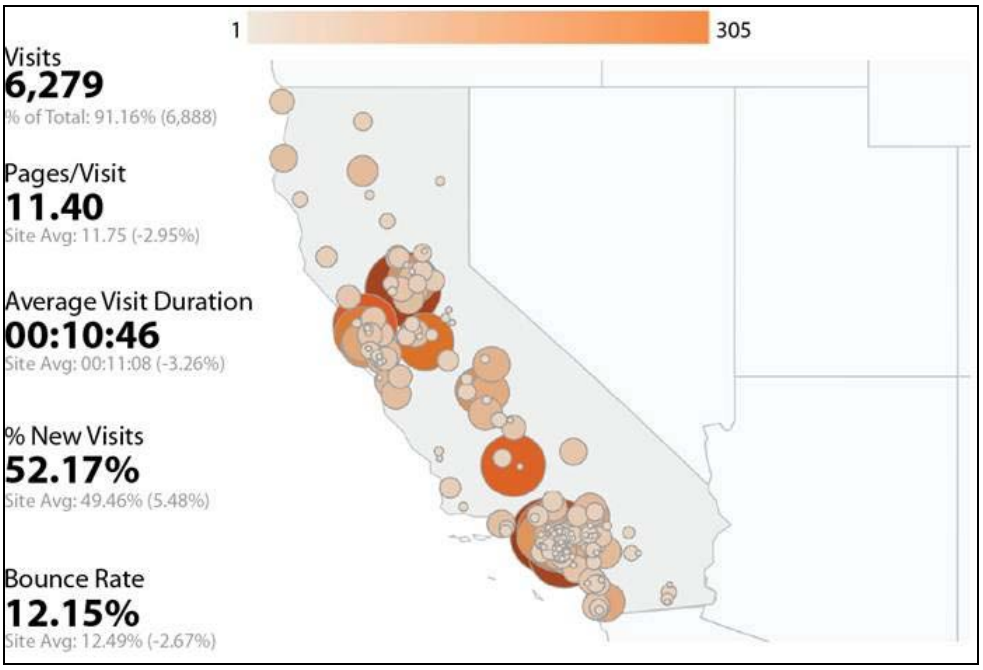


Figure 8. Overview of Management Tools Visitors

Analyses of Adequacy of Customer Support

STAR Technical Assistance Center

Overall, STAR TAC interactions with school and district staff were positive. There was a high degree of enthusiasm in the field for trying out CBT. When callers had technical issues, support was provided successfully. The most common issues during the CBT Tryout reported were the following:

- Confusion with the optional "Location" and "Session" fields in the Pre-ID layout that were used to group test sessions to allow for easier ticket printing. The confusion was caused by the student tickets, which also referred to a Session ID that was not related to the grouping session. Additionally, the alpha- and numeric-only requirement for the "Location" and "Session" fields, respectively, was a challenge for many people.
- Administrators reported that many students thought the "Continue button" was an extension of Part 1 and inadvertently went on to Part 2 instead of stopping at the end of Part 1 as required.
- At times, when a student returned to a test that was paused, that student could not resume testing even though the allotted period for pausing had not reached the time limit.
- If a school-level ID was used to load the Pre-ID, the District and Super-user access levels were not able to see that loaded file without logging on through the use of a school-level ID and password.

Results on Artificial Intelligence Scoring

This section present results of the artificial intelligence scoring systems, m-rater and c-rater, that were used in this study.

M-rater Results

For the seven constructed-response items scored using m-rater, four items were numeric entry that required the student to enter a numeric response, and three items were graph responses that required the student to plot specific points on a grid. As described earlier, all items were scored as right or wrong. For the graphing items, either snap-to-grid was employed or allowances were made to permit a correct answer for those responses that were in close proximity rather than the exact location of the correct answer to account for possible imprecise plotting by students.

As described earlier, to evaluate m-rater results, the m-rater results were compared to the results from a scoring methodology that used an alternate computer scoring mechanism. Complete agreement in awarding and denying points to the students' responses was noted (see Table 26).

The item responses scored by m-rater are described by type—either numeric entry (N.E.) or graph. The number of students with correct, incorrect, or omitted responses are presented, as well as the proportion of students receiving a correct score (***p*-value**) and the point-biserial correlation (**pt-bis**) between the item and the total score. Some items were quite difficult as evidenced by the item statistics. For example, item CSZB0002-01 was a very difficult item (*p*-value=0.01; pt-bis=0.04). This is a graph item that required students to plot three pairs of Cartesian coordinates on a grid, in a piecewise linear manner. In comparison, item CSZB0033, another graph requiring students to plot six pairs of Cartesian coordinates, all in a straight line, was much easier (*p*-value=0.71; pt-bis=0.45). Another difficult item, CSZB0021-04 (*p*-value=0.03, pt-bis=0.18), is a numeric entry item that required students to enter a number indicative of the net force acting on a 0.50-kg car while moving at a constant speed. Note, the two graph items with the largest number of omissions were administered at the end of the first test part.

Table 26. CBT Tryout Items Scored with M-rater

Item	Type	Grade	Section	Position in Section	Correct	Incorrect	Omit	p-value	pt-bis
CSBB0001.03	N.E.	Biology	2	3	1,198	3,870	77 (1%)	0.23	0.30
CSBB0001.07	N.E.	Biology	2	7	860	4,161	124 (2%)	0.17	0.15
CSZB0002.01	Graph	5	2	1	83	8,939	62 (1%)	0.01	0.04
CSZB0021.02	N.E.	8	2	2	1,168	4,477	68 (1%)	0.20	0.51
CSZB0021.04	N.E.	8	2	4	199	5,445	69 (1%)	0.03	0.18
CSZB0032	Graph	8	1	28	1,028	4,745	238 (4%)	0.17	0.45
CSZB0033	Graph	8	1	30	4,296	1,511	204 (3%)	0.71	0.45

C-rater Results

Six of the seven constructed-response items scored with c-rater also had a subset of the responses scored using two human scorers in order to "train" the scoring models. The seventh item, CSBB0001.02, used an alternate computer program to perform the annotations for scoring, since the answers were a combination of capitalized and noncapitalized letter pairs representing Punnett square results. Experience suggests that such items are best scored by computer as humans tend to make scoring mistakes given the subtle differences in responses.

The results of human scoring for the six human-scored c-rater items are presented in Table 27. For each item, the number of responses collected to train the scoring model **(No.)** is presented along with the mean (**Mean**) and standard deviation (**SD**) of each scorer, the standardized mean score difference (**Std Diff**) between the two raters, the quadratic-weighted kappa (**Kappa**), the percentage of exact (**% Agree**) and adjacent agreement (**% Adj Agree**), as well as the correlation between the two sets of ratings. The quadratic-weighted kappa disproportionately weights widely divergent scores as negative; the quadratic-weighted kappa is similar to a correlation since both are based on discrete scores and have the same marginal distribution. As shown in Table 27, the human raters were very much in agreement when scoring the short-text questions.

Table 27. Double Human Scoring of the CBT Tryout Short-text Items Scored

		Human 1		Human 2		Agreement Statistics				
Item	No.	Mean	SD	Mean	SD	Std Diff	Kappa	% Agree	% Adj Agree	Corr
CSBB0001-01	250	0.56	0.50	0.55	0.50	–0.02	0.96	98.0	100	0.96
CSBB0001-06	250	0.60	0.49	0.56	0.50	–0.07	0.93	96.8	100	0.94
CSZB0002-02	250	0.41	0.49	0.41	0.49	0.00	1.00	100.0	100	1.00
CSZB0002-03	250	0.83	0.38	0.82	0.38	–0.01	0.99	99.6	100	0.99
CSZB0006	250	0.16	0.36	0.13	0.34	–0.08	0.82	95.6	100	0.83
CSZB0021-05	250	0.23	0.42	0.23	0.42	0.00	1.00	100.0	100	1.00

The degree of agreement between the first human scorer and c-rater is listed in Table 28. The c-rater-to-human agreement is slightly lower (i.e., degraded) than the human-to-human agreement, but within acceptable limits. Such degradation is frequently observed when moving from scoring solely by humans to automated scoring because the nuances of language can influence how humans view a response compared to the capabilities of a machine. Comparing Table 27 and Table 28, degradation is noted, as indicated by the kappa values and the correlation coefficients.

Table 28. Results of C-rater Agreement with Human Scoring

		Human 1		C-rater*		Agreement Statistics				
Item	No.	Mean	SD	Mean	SD	Std Diff	Kappa	% Agree	% Adj Agree	Corr
CSBB0001-01	250	0.56	0.50	0.52	0.50	–0.08	0.90	95.2	100	0.91
CSBB0001-06	250	0.60	0.49	0.56	0.50	–0.07	0.93	96.4	100	0.93
CSZB0002-02	250	0.41	0.49	0.43	0.50	0.03	0.97	98.4	100	0.97
CSZB0002-03	250	0.83	0.38	0.83	0.38	0.00	1.00	100.0	100	1.00
CSZB0006	250	0.16	0.36	0.15	0.36	–0.01	0.74	93.2	100	0.74
CSZB0021-05	250	0.23	0.42	0.23	0.42	0.00	1.00	100.0	100	1.00

Note: Summary statistics were computed using c-rater scores that were rounded to integers.

The results for the items scored with c-rater after models were built are presented in Table 29. For each item, the number of students with correct, incorrect, or omitted responses as scored by c-rater is listed, along with the proportion of students that received a correct score (***p*-value**) for each task, and the point-biserial correlation (**pt-bis**) between the item and the total score.

Table 29. Results of C-rater–Scored Short-text Constructed-response Items

Item	Type	Grade	Part	Position in Part	Correct	Incorrect	Omit	p-value	pt-bis
CSBB0001.01	c-rater	Biology	2	1	2,531	2,524	90 (2%)	0.49	0.53
CSBB0001.02	c-rater	Biology	2	2	1,004	4,064	77 (1%)	0.20	0.46
CSBB0001.06	c-rater	Biology	2	6	2,864	2,164	117 (2%)	0.56	0.62
CSZB0002.02	c-rater	5	2	2	3,782	5,234	68 (1%)	0.42	0.41
CSZB0002.03	c-rater	5	2	3	7,251	1,764	69 (1%)	0.80	0.42
CSZB0006	c-rater	5	1	14	1,318	8,190	140 (1%)	0.14	0.25
CSZB0021.05	c-rater	8	2	4	1,421	4,213	79 (1%)	0.25	0.55

Comparing the results from m-rater items and c-rater items in Table 26 and Table 29, the percentage of omitted responses for N.E. and graph items (i.e., about two percent on average) is slightly higher compared to the constructed-response items (i.e., about one percent on average). Some of that difference may have been due to the placement of two graph items at the end of the one part. Research (Bridgeman, 1992) shows that constructed-response items involving mathematics computation tend to be more difficult than those same questions with answer choices. Regardless, the number of omitted responses to both sets of constructed-response items is smaller than five percent, a typical flagging criteria used to identify items with high rates of omission.

Conclusion on Test Administration and Artificial Intelligence Scoring

This section of the report presented summary information on LEA, school, and student participation, including analyses of sample representativeness. While many selected schools were unable to take part, the participants were representative of the demographic groups targeted in the sampling plan. The post-test survey responses received from administrators also provided details on what did and did not work well for participating schools and LEAs and how prepared schools and LEAs are for any future testing on CBT. While caution should be taken before drawing general conclusions about California's preparedness for CBT based on the results of the CBT Tryout, the findings in this report may be a useful supplement to other state efforts to promote a smooth transition to next-generation assessments in California. In addition to the results in this report, the *Tips for Computer-Based Testing Transitions in California* (see Appendix D on page 108) were developed based on these survey results and comments from participating educators.

Also presented in this section was an evaluation of the artificial intelligence scoring engines, m-rater and c-rater. Both scoring engines performed well within published quality-scoring guidelines (Williamson, Xi, & Breyer, 2012).

Psychometric Studies

Purpose of the Psychometric Study

The CBT Tryout provided an opportunity to use the data to begin to understand some important issues related to the introduction of CBT and TE items for large-scale assessments. Psychometric analyses were performed to investigate the following research questions:

1. What are the statistical properties of the CBT Tryout forms and items comprising the forms?
2. Do TE items measure the same underlying construct as traditional MC items for each of the three tests?
3. Does CBT lead to differential impact on student performance at the item and test level for subgroups of interest classified by key demographic variables (e.g., ethnicity, socioeconomic status [SES] or parent/guardian level of education)?
4. What are the major factors that impact students and school readiness for CBT? How do these factors contribute to students' performance on the CBT Tryout?

Data and methods used to address these research questions as well as the analyses results are described in each of the following sections:

1. Test and Item Analyses
2. Dimensionality of the CBT Tryout Assessments
3. Differential Impact of CBT on Student Performance
4. Student- and School-Level Readiness Factors for CBT Implementation

Test and Item Analyses

This section of the report presents information regarding the statistical properties of the CBT Tryout forms and items comprising the forms, for each grade, including:

1. Total test summary statistics, reliability coefficients, and raw score distributions;
2. Summary statistics for TE items;
3. Summary statistics for MC items, from both the CBT and the original pencil-and-paper (PPT) testing environment;
4. Scatterplots of CBT and PPT MC item p-values; and
5. Omit rates.

Data

As noted earlier, because the CBT Tryout occurred in the fall, test forms were administered to students who had completed the content to be tested in the preceding academic year—that is, the grade five form was administered to grade six students, the grade eight form was administered to grade nine students, and the Biology form was administered to students who had completed a Biology course in the previous academic year. As noted in Table 2, the total number of students tested included 9,659 students for the grade five science test; 6,032 students for the grade eight science test; and 5,404 students for the Biology test.

Not-presented Responses and Omitted Responses

As described earlier in the *Discussion of the Missing and Incomplete Data* section of this report, a number of student records had an invalid start time or end time (i.e., technology failure, not ending the test properly, or student was absent). Although blank responses are typically treated as omitted in tests with untimed sessions, in the CBT Tryout, blank item responses associated with invalid start or end times were coded as "not presented." That is, after the last valid response, all blank responses were coded as "not presented" and were not scored. Similarly, if the entire part was left blank, all responses in that part were also coded as "not presented." However, if the record had a valid start time and end time and the student simply left some items blank but had valid responses in other items, the blank responses were coded as "missing" and were given a score of zero.

All student records with items responses marked as "not presented" were treated as incomplete records, not scored, and not included in analyses requiring valid total test scores (e.g., point-biserial, reliability, descriptive statistics). However, because students may have attempted some items and made valid responses prior to the not-presented items, the responses to all preceding items were treated as valid responses and included in the computation of the associated *p*-value and omit rates for those items.

Methods

Test-level and item-level statistics were calculated using SAS version 9.2, a standard statistical analysis software package (SAS Institute Inc., 2009). Test-level analyses for each test included mean and standard deviation (SD) of the total raw score, reliability coefficient as measured by Cronbach's alpha (Cronbach, 1951), and raw score at the 10^{th}, 25^{th}, 50^{th}, 75^{th} and 90^{th} percentiles.

Item statistics included those statistics typically described as "classical item analyses" and included the percentage of examinees in the sample that answered the item correctly (*p*-value), point-biserial

correlation between item and performance of the total test (pt-bis), and percentage of students omitting each item (percent omit).

Results

Summary Statistics for Total Test

The total test raw score summary statistics, reliability coefficient, mean *p*-value, mean point-biserial (pt-bis) correlation, and raw score at the 10th, 25th, 50th, 75th and 90th percentiles, for the test at each grade are presented in Table 30. As noted in the table, the maximum obtainable score of 60 was observed in grade eight and Biology; in grade five, the maximum observed score was only 56. The reliability coefficients are relatively high across the three tests, ranging from 0.88 to 0.91. Raw score distributions for each of the three tests are presented in Figure 9 through Figure 11.

Table 30. CBT Raw Score Summary Statistics

	No. Students	No. Items *	Mean	SD	Min Obtained	Max Obtained	Mean *p*-value	Mean pt-bis	Reliability	P10	P25	P50	P75	P90
Grade 5	9,018	60	30.4	9.4	0	56	0.51	0.34	0.88	18	23	30	37	43
Grade 8	5,648	60	32.2	11.1	0	60	0.53	0.41	0.91	18	24	32	40	47
Biology	5,051	60	29.2	10.9	0	60	0.49	0.38	0.90	15	20	29	37	44

* Each item contributes 1 point to the total raw score. The possible score range is 0–60 for all tests.

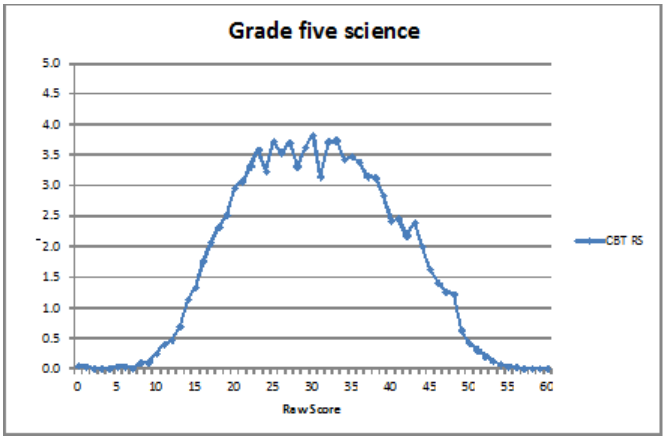


Figure 9. Raw Score Distribution for Grade Five Science

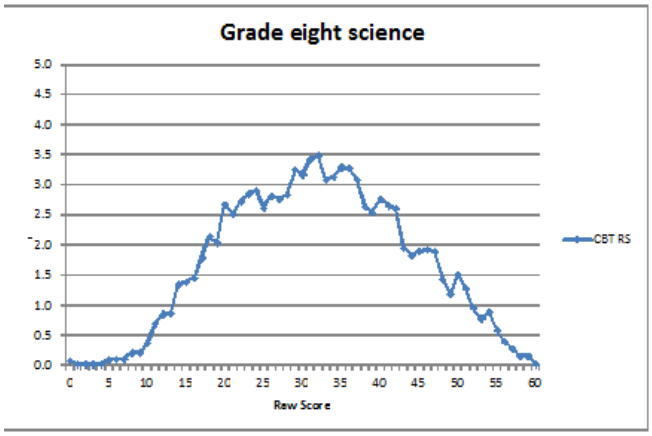


Figure 10. Raw Score Distribution for Grade Eight Science

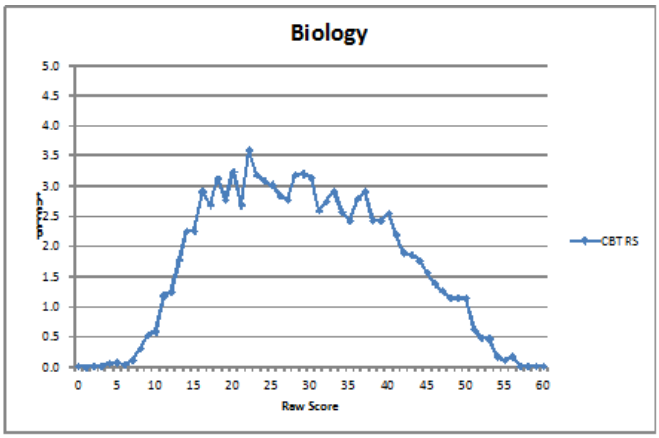


Figure 11. Raw Score Distribution for Biology

Summary Statistics for TE and MC Items

Summary statistics for the TE items, including *p*-values, point-biserial correlations, and percent omits are included in Table 31.

Table 31. TE Item Summary Statistics for Each Test

	No. Items	*p*-value		Point-Biserial		Percent Omit	
		Mean	SD	Mean	SD	Mean	SD
Grade 5	20	0.40	0.30	0.28	0.12	1.71	2.47
Grade 8	20	0.43	0.23	0.46	0.08	2.08	1.23
Biology	20	0.35	0.16	0.39	0.14	1.38	0.95

Parallel information for the multiple-choice items is presented in Table 32. In this table, the information includes the values based on the CBT Tryout administration and, for reference, the stored item bank values for each item. That is, for each of the tests, the first row contains the statistics based on the data from the CBT Tryout and second row contains the statistics based on the most recent spring paper-and-pencil administration, which is stored in the item bank. Due to potential differences in populations over time, caution should be used when making comparisons between the PPT and CBT versions of the summary statistics. The number of students with valid responses, *p*-value, point-biserial correlation, and percent omit for each individual item are located in Appendix E, Table E.1 through Table E.3, starting on page 110. Note that calculations of the *p*-value and omit rate of each item were based on all students who responded to the item after excluding cases where the item was coded as "not presented." However, because the point-biserial is calculated as an item-total correlation, calculations were based on students with valid total scores only, specifically those who responded to all items after excluding any cases with a "not presented" for any test item.

Table 32. Reused (MC) Item Summary Statistics for Each Test

		No. Items	*p*-value		pt-bis		Percent Omit	
			Mean	SD	Mean	SD	Mean	SD
Grade 5	CBT	40	0.56	0.14	0.37	0.08	1.10	0.65
	PPT	40	0.54	0.14	0.37	0.08	0.14	0.06
Grade 8	CBT	40	0.59	0.17	0.39	0.08	1.89	0.93
	PPT	40	0.55	0.17	0.38	0.09	0.13	0.05
Biology	CBT	40	0.55	0.13	0.38	0.08	0.93	0.63
	PPT	40	0.56	0.11	0.42	0.08	0.14	0.05

Scatterplots of CBT and PPT MC Items

Figure 12 through Figure 14 are scatterplots comparing the *p*-values of the CBT and PPT modes of administration for the MC items for each test. The scatterplots demonstrate that very few items fall beyond the measurement error band of the regression line (the broken lines in each plot) between PPT and CBT *p*-values. The correlations between MC item *p*-values from the CBT and PPT administrations are 0.92 for each test. Although the correlations between MC item point-biserials from the two administration modes are lower, ranging from 0.70 to 0.75 (see Appendix E, Table E.1 through Table E.3, starting on page 110), together these results indicate that the items are functioning similarly between CBT and PPT administrations with respect to item difficulty and item discrimination as measured by the point-biserial values. As mentioned in the previous section, statistics for the PPT version of the MC items are item bank values obtained from the spring 2012 CST administration. As such, caution should be used when generalizing the comparison of item statistics between PPT and CBT version of the same item, due to potential differences in populations.

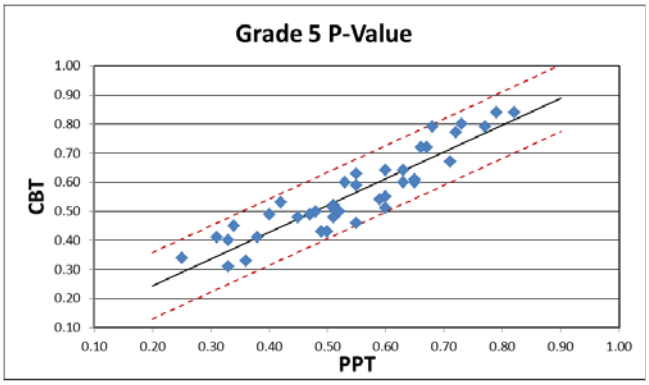


Figure 12. CBT vs. PPT *p*-values, Grade Five Science Reused (MC) Items

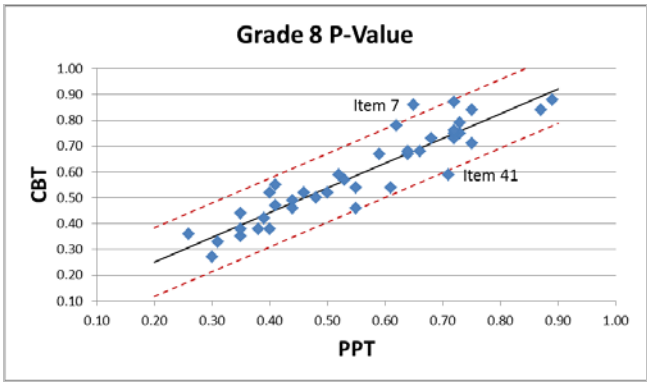


Figure 13. CBT vs. PPT *p*-values, Grade Eight Science Reused (MC) Items

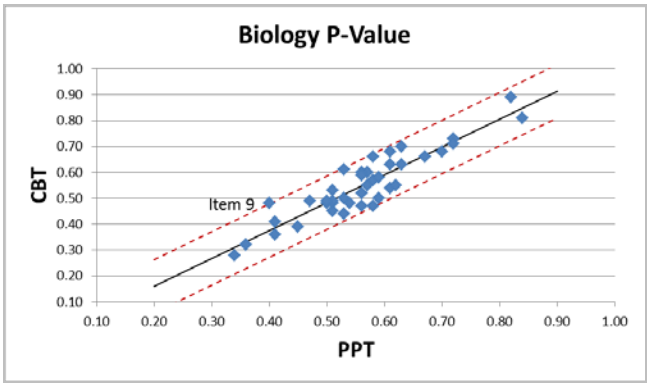


Figure 14. CBT vs. PPT *p*-values, Biology Reused (MC) Items

Omit Rates

As shown in Appendix E (Table E.1 through Table E.3), there were some differences, albeit relatively minor, in omit rates between administration mode and item type. The omit rates for MC items administered in PPT appear to be similar regardless of the position of the item in the original test, whereas in the CBT Tryout, the omit rate tends to increase toward the end of each test part. In addition, the overall omit rates for the MC items administered in the CBT Tryout, while relatively low, were slightly higher than when the same MC items were administered on PPT forms. These variations

in omit rates may be explained by a lack of motivation of students due to the low-risk, no-accountability nature of these CBT Tryout test forms.

In addition, the omit rate for some of the TE items was marginally higher than the omit rates observed for the MC items located immediately before or after the TE items, or some combination of these factors. This finding may be attributed to a combination of low student motivation and the relatively higher cognitive demand of the TE items. This finding may appear somewhat contradictory to the post-administration survey results indicating that students found the CBT and TE items more engaging than the traditional MC items administered on paper. However, this potential discrepancy may be due to the fact that the post-administration survey results are subjective measurements based on student perceptions, rather than student performance.

Of all the items in the CBT Tryout forms, there was one item with an omit rate greater than five percent: grade five science item #45 had an omit rate of 11.78. This was a TE item designed to better address a standard requiring students to demonstrate knowledge and understanding by constructing a simple circuit. The item was designed to allow students to demonstrate their knowledge and skills through interacting with moveable objects. Although the item is a better match to the expectations of the standard than a traditional PPT item, it is the first time students have been asked to demonstrate such knowledge and skills in this format. Furthermore, the item is both highly challenging from a cognitive perspective and is measuring a very difficult construct, which may explain the high omit rate.

In addition, omit rates were calculated for each item based on students' performance level classifications on their spring 2012 CST science tests (see Table E.7 through Table E.9 in Appendix E starting on page 119). The average omit rate for the TE and MC items is presented for each performance level by test form in Table 33.

Across performance levels for all forms, omit rates for TE items were higher on average than omit rates for MC items. In addition, the average omit rates for TE items tend to decrease as the proficiency level increases for all three tests, with the exception of the far below basic and below basic levels of the Biology test. A similar trend holds for omit rates for MC items on the grade eight test, but not for the grade five and Biology tests. For the grade five test, students in the proficient level had the highest average omit rate for MC items (1.20 percent) among the five performance levels. For Biology, students who were far below basic and below basic had lower average omit rates on MC items than students who were in the basic or proficient performance levels.

Table 33. CBT Tryout Percent Omit Summary Statistics by Performance Level on the Spring 2012 CST for TE and MC Items

			TE Percent Omit		MC Percent Omit	
		N	Mean	SD	Mean	SD
Grade 5	Far below basic	495	2.15	4.63	0.92	0.87
	Below basic	750	2.03	4.00	1.05	0.62
	Basic	2,215	1.73	2.99	0.95	0.64
	Proficient	3,207	1.67	2.11	1.20	0.69
	Advanced	2,347	1.14	1.23	0.88	0.63
Grade 8	Far below basic	328	2.52	1.46	2.31	1.26
	Below basic	396	2.47	1.17	1.90	1.03
	Basic	751	2.20	1.64	1.86	1.15
	Proficient	1,070	1.87	1.11	1.58	0.85
	Advanced	2,592	1.26	0.98	1.03	0.49

			TE Percent Omit		MC Percent Omit	
		N	Mean	SD	Mean	SD
Biology	Far below basic	272	1.08	1.93	0.40	0.34
	Below basic	370	1.33	1.21	0.57	0.45
	Basic	1,456	1.50	0.97	1.04	0.48
	Proficient	1,508	1.38	1.05	0.73	0.50
	Advanced	1,305	0.88	0.65	0.57	0.48

Omit rates were also calculated for subgroups of interest by test: English learners vs. non–English learners and economically disadvantaged vs. non–economically disadvantaged. Tables in Appendix E (Table E.4 through Table E.6, starting on page Table E.4. CBT Percent Omit for Grade Five by EL Status and Economic Status115) list the percent-omit information for each item for each of the three tests. The omit rates presented in Table 34 are the average omit rate by test form for the TE and MC items. Similar to the overall groups, there were some small difference in omit rates between TE and MC items. With the exception of MC items on the Biology test, English learners tended to have higher rates of omission on both TE and MC items in comparison to non–English learners. Results for economically disadvantaged vs non–economically disadvantaged groups varied across tests and item types. All average omit rates were less than three percent.

Table 34. CBT Tryout Percent Omit Summary Statistics by English Proficiency and Economic Status on the Spring 2012 CST for TE and MC Items

			Technology-Enhanced Percent Omit		Multiple-Choice Percent Omit	
		N	Mean	SD	Mean	SD
Grade 5	English learner	2,352	2.01	3.32	1.28	0.78
	Non-English learner	7,019	1.55	2.17	0.99	0.61
	Economically disadvantaged	5,935	1.66	2.66	1.04	0.69
	Not economically disadvantaged	3,433	1.67	2.11	1.10	0.59
Grade 8	English learner	773	2.80	1.56	2.56	1.24
	Non-English learner	4,577	1.61	1.07	1.28	0.62
	Economically disadvantaged	3,335	1.94	1.25	1.60	0.79
	Not economically disadvantaged	2,015	1.52	0.95	1.24	0.59
Biology	English learner	554	1.60	1.29	0.76	0.50
	Non-English learner	4,720	1.34	0.92	0.91	0.64
	Economically disadvantaged	3,061	1.37	0.93	0.84	0.50
	Not economically disadvantaged	2,241	1.36	1.00	0.98	0.80

Dimensionality of the CBT Tryout Assessment

This section of the report addresses the research question: Do TE items measure the same underlying construct as traditional MC items for each of the three tests? This question was addressed by examining the dimensionality of the CBT Tryout data. The primary goal was to gain some understanding as to whether each of the CBT Tryout assessments is represented by a single factor or dimension.

Data

The data used to examine the dimensionality included all data with the exception of the incomplete records described earlier. For this set of analyses, the data for each test were randomly partitioned into an exploratory (EFA) subsample and a confirmatory (CFA) subsample as noted in Table 35.

Table 35. Dimensionality Analysis Sample Sizes by Science Test

	Total	EFA	CFA
Grade 5	9,018	4,546	4,472
Grade 8	5,648	2,817	2,831
Biology	5,051	2,576	2,475

Method

Exploratory factor analysis (EFA) and confirmatory factor analysis (CFA) were used to examine the extent to which each CBT Tryout test supports a unidimensional interpretation. An adequate data fit to a unidimensional model would provide evidence that the TE items are measuring the same underlying construct as the traditional MC items. All of the exploratory and confirmatory analyses were conducted using Mplus (Muthen & Muthen, 2009), a statistical modeling software program employed for complex data analyses.

Exploratory Factor Analyses

The dimensionality of each CBT Tryout was first explored by means of EFA, using the subsamples designated for these analyses. The factor loadings were extracted using the maximum likelihood estimation with a promax rotation. Iterative solutions were tested for one to four factors for each test. Each solution was evaluated for its ability to produce dimensions that: (a) satisfy Cattell's (1966) scree test; (b) retain two or more items with salient loadings, where loadings > 0.25 are considered salient; (c) yield reasonable internal consistency for items with salient loadings; (d) have eigenvalues greater than 1.0 as specified by the Kaiser-Guttman rule (Brown, 2006); and (e) make psychological sense in terms of parsimonious solutions including mutually exclusive assignment of items to factors, maximum number of items retained, and compatibility with dimensions obtained in other empirical work (Fabrigar, Wegener, MacCallum, & Strahan, 1999).

Confirmatory Factor Analyses

The CFAs were conducted to compare a unidimensional model and a bifactor model (Gibbons & Hedeker, 1992), using the confirmatory subsamples from each test. The models were compared to determine whether the CBT Tryout assessments introduced any secondary factors for the TE and MC items respectively, and whether a bifactor model fit the data significantly better than a single-factor solution.

The MPlus software enables the evaluation of both models based on Akaike's (1987) information criterion (AIC), Bayesian information criterion (BIC) (Schwarz, 1978) and sample-size adjusted

Bayesian information criterion (SABIC) (Enders & Tofighi, 2008). All these criteria are measures of the relative goodness-of-fit of a statistical model. In general, when comparing these fit statistics for different models smaller values are better, but there is no criterion for significant differences. The fit statistics provide a good starting point for deciding on a model; however, the substantive interpretation of the model coefficients—as they relate to practical significance—should also be considered.

Results

Exploratory Factor Analyses

The scree plots from the EFA for each CBT Tryout test are presented in Figure 15 through Figure 17. For all tests, the results suggested that the first extracted factor was dominant with an eigenvalue much larger than the other factors. This finding indicates a sizeable differentiation between the variance accounted for by the first and second components. Using the Biology test as an example (Figure 17), the ratio of the largest eigenvalue to the next largest value was 8.22. The one-factor model seemed to best fit the data for all three tests and satisfied the evaluation criteria; this finding is consistent with the unidimensional structures observed for results obtained from the PPT California Standards Tests (CSTs) (Gaffeny & Perryman, 2009).

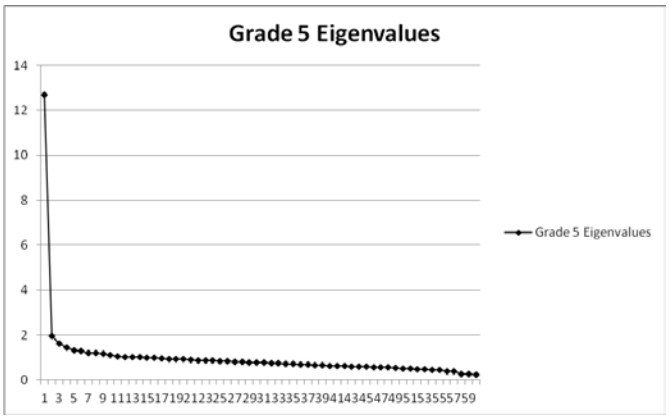


Figure 15. Scree Plot for the Grade Five Science CBT

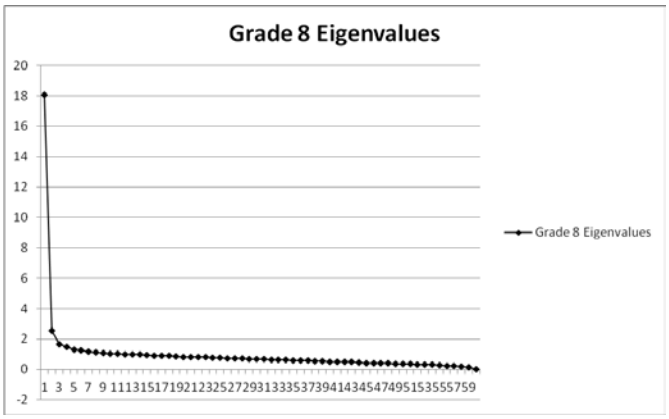


Figure 16. Scree Plot for the Grade Eight Science CBT

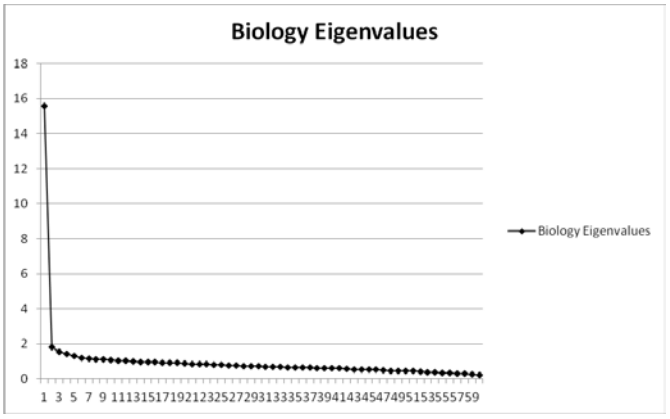


Figure 17. Scree Plot for the Biology CBT

Confirmatory Factor Analyses

The model-fit statistics for all three tests are listed in Table 36. The model-fit statistics for the comparison between the unidimensional model and the bifactor model show that AIC, BIC and SABIC of the bifactor models are lower than those of the unidimensional models except for BIC of the grade five science test. This suggests that the bifactor model fits better than the unidimensional model. However, an examination of the loadings for the secondary factors in the bifactor model suggests that very few items load substantively on these factors at each grade level. That is, the secondary factors appear to be minor dimensions. The better fit of the bifactor models is likely a result of the items that do load on the secondary factors, but from a practical perspective it may not make sense to retain these factors. These results suggest that the data should be treated as essentially unidimensional.

Table 36. Fit Statistics for the Unidimensional Model and Bifactor Model

	Model	AIC	BIC	SABIC
Grade 5	Unidimensional	291,699	292,468	292,086
	Bifactor	291,392	292,545	291,973
Grade 8	Unidimensional	181,122	181,835	181,454
	Bifactor	180,630	181,701	181,129
Biology	Unidimensional	170,887	171,585	171,204
	Bifactor	170,362	171,409	170,837

Following selection of the unidimensional model, model fit was further evaluated using the fit indices that measure the proportionate improvement in fit by comparing the hypothesized model with the baseline model (independence model). These fit guidelines include the comparative fit index (CFI), the non-normed fit index (NNFI, also known as Tucker-Lewis Index, TLI), and the root mean square error of approximation (RMSEA). The CFI ranges from 0 (poor fit) to 1 (good fit). The NNFI is interpreted like the CFI. While there is no single set of evaluation rules agreed upon by all researchers, Hu and Bentler (1999) provide the following guidelines for evaluating fit indices: CFI $\geq$ 0.95, NNFI $\geq$ 0.95, and RMSEA $<$ 0.06. The fit statistics, as shown in Table 37, suggest good model-data fit for the unidimensional model for each of the three confirmatory subsamples.

Table 37. Fit Indices for the Unidimensional Model for Three Tests

	CFI	NNFI	RMSEA
Grade 5	0.98	0.98	0.01
Grade 8	0.96	0.95	0.03
Biology	0.97	0.97	0.02

Differential Impact of CBT on Student Performance

The data obtained from the CBT Tryout provided an opportunity to examine whether the introduction of CBT and TE items may result in differential impact in student performance, at the item level and at the test level. This section of the report addresses the research question: Does CBT lead to differential impact on student performance at the item and test level for subgroups of interest classified by key demographic variables? Demographic groups were defined according to gender, race/ethnicity, English speakers/English learners, no special services/special services, SES, parental education, and area of residency (i.e., metropolitan/small town and rural). Potential group differences were examined at the item level using differential item functioning (DIF) analyses and at the total test level by means of regression and residual analyses.

Data

The data used to address this research question included all data with the exception of the incomplete records, described earlier. Students with complete records were then matched to their records from the 2012 spring operational administration of the CSTs. Specifically, students administered the grade five science CBT forms were matched to their 2012 grade five science CST results; students administered the grade eight science CBT forms were matched to their 2012 grade eight science CST results; and students administered the Biology CBT forms were matched to their 2012 Biology CST results. Summary statistics for all of the groups of interest are presented in Appendix E, Table E.10 through Table E.12, starting on page 123.

Method

Differential Impact at the Item Level: Differential Item Functioning (DIF)

DIF analyses were conducted to investigate potential performance differences among the subgroups of interest that participated in the CBT. These results were then compared to determine if results changed given the mode of administration (CBT or PPT). For the DIF analysis, the male, white, English only, no special education services, urban resident, high SES, and high parent/guardian education level groups were treated as the reference groups; the female, other race and ethnicity, English learner, special education services, small town/rural resident, low SES, and low parent/guardian education groups were considered the focal groups. For the DIF analyses of the CBT items, students' CST scores were used as criterion scores. This provided a more stable measure of the students' science ability and it provided similar basis for comparing the CBT and PPT results. For comparison, the most recent DIF results for the PPT-format MC items were obtained from the CST item bank. Of note, unlike the CBT, which were collected during a single administration, data from the PPT version of these items were collected from various administrations.

DIF statistics are used to identify items whereby identifiable groups of students with the same underlying level of ability (e.g., females, African Americans, English learners) have different probabilities of answering correctly. If the item is more difficult for an identifiable subgroup, the item may be measuring something different from the intended construct. However, it is important to recognize that DIF-flagged items might also be related to actual differences in relevant knowledge or skill (item impact) or related to statistical Type I error. In this study, it is also possible that the DIF results would differ between the CBT and PPT modes of administration because they were administered at different times of the year and characterized by different levels of motivation, since the CBT Tryout was not a high-stakes test. Using students' spring CST scores as criterion scores also presupposed that students' academic abilities remained consistent throughout the year.

The DIF analyses utilized the Mantel-Haenszel (MH) DIF statistic (Holland & Thayer, 1985; Mantel & Haenszel, 1959). This statistic shows the difference in performance on an item between the focal and reference groups after conditioning on total test score. Based on the magnitude of the DIF statistics, items are classified into one of three categories: A, B, or C. Category A contains items with negligible DIF. Category B includes items with slight to moderate DIF. Category C contains items with moderate to large values of DIF. Negative DIF statistics favor the reference group while positive values favor the focal group.

Differential Impact at the Overall Test Level: Regression and Residual Analyses

To investigate the possible differential impact of CBT on the overall performance of students from different demographic groups, a regression of the CBT Tryout scores with the spring 2012 CST science scores was conducted. The regression equation is:

$$Y = \beta_0 + \beta_1 X + \varepsilon$$

where,

Y = CBT score;

β_0 = intercept of the regression,

β_1 = slope of the regression,

X = CST science spring 2012 score, and

ε = residual.

The distribution of residuals was examined for the overall group as well as the student subgroups defined by demographic characteristics (e.g., gender, ethnicity, English fluency, special education status, SES, parent/guardian education, and residency). Similar to DIF analyses, this approach is based on the assumption that a consistent relationship between CST and CBT test scores should hold for all students if differential impact due to CBT is not observed; an additional assumption is that the impact of the time lag between spring and fall testing is consistent across students groups.

Regression residuals are defined as the predicted value subtracted from the observed value. For the overall group, the residuals are expected to have a mean of zero and error variance best estimated by mean squared error. If homogeneity of variance is evident and an absence of bias is observed at particular intervals of the predictor score scale for the overall group, then all members of student subgroups, regardless of their performance level, should have the same residual distribution as the overall group. Different residual distributions for specific subgroups, especially when the mean deviates significantly from zero, indicate differential impact of the CBT testing on different subgroups. It is also useful to investigate which group of students has the largest positive residual mean (i.e., performance exceeding prediction or students being advantaged in the CBT mode) and which group has the largest negative residual mean (i.e., performance below prediction or students being disadvantaged in the CBT mode).

Results

Differential Impact at the Item Level

Items flagged for C-DIF from the CBT Tryout are summarized in Table 38, Table 40, and Table 42. The grade five science test had two MC items flagged for C-DIF; one favored white over African American students and the other favored small town/rural over metropolitan students. The grade eight science test had three TE items and three MC items flagged for C-DIF; one TE item favored small town and rural students, one TE item and one MC item favored English speakers, while

one TE item and two MC items favored the white students over Hispanic students and Asian and Combined Asian students, respectively. Five TE items and three MC items were flagged for C-DIF on the Biology test; one MC item favored nondisabled students, and five TE items and two MC items favored students living in metropolitan areas.

Table 39, Table 41, and Table 43 provide the DIF results from the CBT Tryout along with the most recently available DIF results from the PPT administrations for all the MC items that were flagged for B- or C-DIF, in either testing mode. Detailed DIF results for each test are presented in Table E.13 through Table E.18 in Appendix E starting on page 134. None of the items flagged for C-DIF in the CBT Tryout were flagged for C-DIF when they were used in PPT mode. While every effort was made to make the CBT version comparable to the PPT version (i.e., items were carefully reviewed and compared, between the two modes of administration, by content experts to verify that changes from PPT to CBT, if any, were negligible), changes in the DIF category may indicate that computer delivery had some impact on students' performance. It should be noted that PPT operational analyses do not routinely conduct DIF comparisons for disability, SES, area of residency, or parent/guardian education; therefore the comparison between the CBT and PPT results for these subgroups was not possible.

Table 38. C-DIF Items for Grade Five Science

Item	Reference/Focal DIF Comparison	CBT	
		Cat.	MH D-DIF
5 (MC)	White/African American	C-	–2.14
7 (MC)	Metropolitan/Small town/Rural	C+	1.62

Table 39. Grade Five Science: MC Items Flagged for B- or C-DIF from Either Test Mode

Item	Reference/Focal DIF Comparison	CBT		PPT	
		Cat.	MH D-DIF	Cat.	MH D-DIF
2 (MC)	White/Asian American	A-	–0.61	B-	–1.02
2 (MC)	White/Filipino	A-	–0.36	B-	–1.08
2 (MC)	White/Hispanic	B-	–1.24	A-	–0.64
2 (MC)	White/Combined Asian	A-	–0.50	B-	–1.01
5 (MC)	Male/Female	B-	–1.11	A-	–0.69
5 (MC)	White/Filipino	A-	–0.76	B-	–1.09
5 (MC)	White/African American	C-	-2.14	B-	–1.02
6 (MC)	White/Hispanic	B-	–1.06	A-	–0.53
7 (MC)	Male/Female	B-	–1.17	A-	–0.97
7 (MC)	White/Hispanic	B-	–1.07	A-	–0.56
25 (MC)	White/Asian American	B+	1.03	A+	0.71
29 (MC)	White/Filipino	B-	–1.05	A-	–0.84
37 (MC)	White/Filipino	B-	–1.15	A+	0.04
50 (MC)	White/African American	B-	–1.07	A-	–0.20
53 (MC)	White/Filipino	B-	–1.05	A-	–0.36
57 (MC)	White/Filipino	B+	1.29	A+	0.33

Table 40. C-DIF items for Grade Eight Science

Item	Reference/Focal DIF Comparison	CBT	
		Cat.	MH D-DIF
19 (TE)	Metropolitan/Small town/Rural	C+	1.76
31 (TE)	English speaker/English learner	C-	–2.00
35 (TE)	White/Hispanic	C-	–1.57
42 (MC)	English speaker/English learner	C-	–1.70
51 (MC)	White/Asian American	C-	–1.58
51 (MC)	White/Combined Asian	C-	–1.56

Table 41. Grade Eight Science: Items Flagged for B- or C-DIF from Either Test Mode

Item	Reference/Focal DIF Comparison	CBT		PPT	
		Cat.	MH D-DIF	Cat.	MH D-DIF
7 (MC)	Male/Female	B+	1.17	B+	1.28
7 (MC)	White/African American	B-	–1.59	A+	0.02
9 (MC)	White/Asian American	B-	–1.12	A-	–0.09
9 (MC)	White/Combined Asian	B-	–1.02	A-	–0.09
21 (MC)	White/Asian American	B-	–1.02	A-	–0.29
21 (MC)	English speaker/English learner	B-	–1.27	A-	–0.18
23 (MC)	White/African American	B-	–1.21	A-	–0.42
27 (MC)	White/Hispanic	B-	–1.43	A-	–0.19
27 (MC)	White/African American	B-	–1.34	A-	–0.16
29 (MC)	White/Filipino	B+	1.87	A+	0.50
37 (MC)	White/Asian American	B-	–1.26	A-	–0.45
37 (MC)	White/African American	B-	–1.58	A+	0.12
37 (MC)	White/Combined Asian	B-	–1.02	A-	–0.40
42 (MC)	Male/Female	B-	–1.00	A-	–0.56
42 (MC)	White/Hispanic	B-	–1.20	A-	–0.72
42 (MC)	English speaker/English learner	C-	–1.70	B-	–1.14
46 (MC)	English speaker/English learner	B-	–1.09	A-	–0.53
50 (MC)	Male/Female	B-	–1.30	A-	–0.90
50 (MC)	White/Asian American	B-	–1.59	A+	0.38
50 (MC)	White/Filipino	B-	–1.95	A+	0.22
50 (MC)	White/Combined Asian	B-	–1.57	A+	0.31
51 (MC)	Male/Female	B-	–1.16	B-	–1.25
51 (MC)	White/Asian American	C-	–1.58	B-	–1.10
51 (MC)	White/Filipino	B-	–1.35	B-	–1.01
51 (MC)	White/Hispanic	B-	–1.01	A-	–0.68
51 (MC)	White/Combined Asian	C-	–1.56	B-	–1.06
52 (MC)	Male/Female	B-	–1.05	A-	–0.96
53 (MC)	Male/Female	B-	–1.50	A-	–0.91
55 (MC)	White/Filipino	B-	–1.07	A-	–0.12
56 (MC)	White/Hispanic	B-	–1.07	A+	0.01
58 (MC)	Male/Female	B+	1.04	B+	1.19
58 (MC)	White/Asian American	A+	0.84	B+	1.13

Table 42. C-DIF Items for Biology

Item	Reference/Focal DIF Comparison	CBT	
		Cat.	MH D-DIF
9 (MC)	No special services/Special services	C-	–1.85
11 (TE)	Metropolitan/Small town/Rural	C-	–1.78
17 (MC)	Metropolitan/Small town/Rural	C-	–1.69
33 (TE)	Metropolitan/Small town/Rural	C-	–2.49
35 (TE)	Metropolitan/Small town/Rural	C-	–1.65
36 (TE)	Metropolitan/Small town/Rural	C-	–2.49
37 (TE)	Metropolitan/Small town/Rural	C-	–1.93
53 (MC)	Metropolitan/Small town/Rural	C-	–1.85

Table 43. Biology: Items Flagged for B- or C-DIF from Either Test Mode

Item	Reference/Focal DIF Comparison	CBT		PPT	
		Cat.	MH D-DIF	Cat.	MH D-DIF
6 (MC)	White/Filipino	B+	1.94	A+	0.17
7 (MC)	White/Filipino	B-	–1.70	A-	–0.12
12 (MC)	English speaker/English learner	B-	–1.03	A-	–0.64
14 (MC)	Male/Female	B+	1.40	A+	0.85
46 (MC)	White/Asian American	B-	–1.05	B-	–1.14
47 (MC)	White/Filipino	B+	1.65	A+	0.55
50 (MC)	White/Asian American	B+	1.41	A-	–0.74
50 (MC)	White/Combined Asian	B+	1.50	A-	–0.59
55 (MC)	White/Filipino	B-	–1.19	A-	–0.41
60 (MC)	White/Asian American	B+	1.56	A+	0.09
60 (MC)	White/Filipino	B+	1.03	A+	0.00
60 (MC)	White/Combined Asian	B+	1.34	A+	0.06

Differential Impact at the Overall Test Level

The scatterplots of CBT Tryout raw scores and 2012 CST scale scores for each test are illustrated in Figure 18 through Figure 20 for each of the three tests. In general, raw score points from the CBT Tryout are spread evenly around the regression line, indicating both linearity and homogeneity of variance between the two sets of test scores. The regression results and fit statistics are presented in Table 44 and Table 45. The coefficients of determination, denoted R-square, ranged from 0.62 to 0.65 across the three tests, indicating that 62 to 65 percent of variance in CBT Tryout scores can be explained by spring 2012 science CST scores.

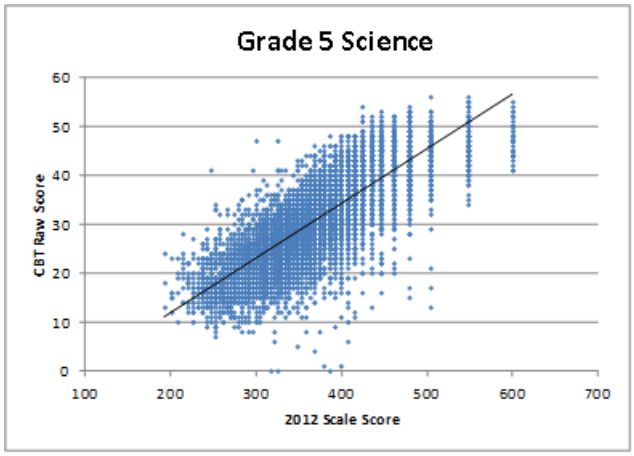


Figure 18. Grade Five Science CBT Raw Score vs. 2012 Scale Score

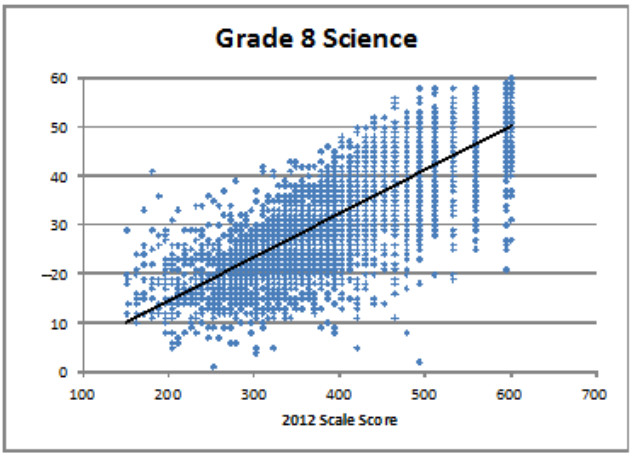


Figure 19. Grade Eight Science CBT Raw Score vs. 2012 Scale Score

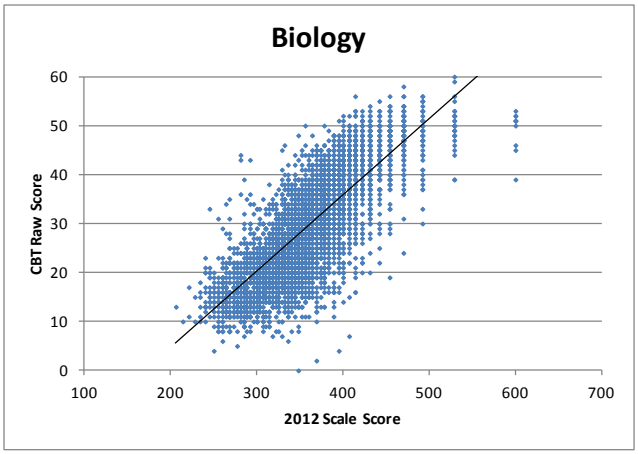


Figure 20. Biology CBT Raw Score vs. 2012 Scale Score

Table 44. Parameter Estimates for Regression of CBT Raw Score on 2012 Scale Score

		Parameter Estimates				
		DF	Estimate	SE	t-value	prob(t)
Grade 5	Intercept	1	–10.52	0.35	–30.37	<.001
	2012 Scale Score	1	0.11	0.00	121.33	<.001
Grade 8	Intercept	1	–3.60	0.39	–9.17	<.001
	2012 Scale Score	1	0.09	0.00	96.01	<.001
Biology	Intercept	1	–26.64	0.66	-40.61	<.001
	2012 Scale Score	1	0.16	0.00	86.66	<.001

Table 45. Fit Statistics for Regression of CBT Raw Score on 2012 Scale Score

	Fit Statistics			
	Dependent Mean	Coeff. Var.	R-Square	Adj. R-Square
Grade 5	30.85	17.99	0.64	0.64
Grade 8	33.10	19.16	0.65	0.65
Biology	29.57	22.39	0.62	0.62

The distributions of residuals were examined for the total group studied as well as the demographic subgroups of interest. Mean residuals for specific subgroups are presented in Table 46 through Table 48. Cohen's *d* (Cohen, 1992) was calculated as effect size indicator for comparison of the mean residuals between subgroups and was presented in Table E.19 through Table E.21 starting on page 136. When interpreting effect sizes, a general rule of thumb regarding the magnitude of Cohen's *d* is as follows: values around 0.2 indicate a small effect, values around 0.5 indicate a moderate effect, and values greater than 0.8 indicate a large effect.

The differences between the mean residuals of the subgroups were small to moderate according to the effect sizes, yet there seems to be a pattern across the three tests. For example, the members of the Asian American and White subgroups had higher residual means than Hispanics and African Americans (effect sizes range from 0.03 to 0.41). This suggests that Asian American and White students performed better than expected on the CBT Tryout in contrast to the Hispanic and African American students when controlling for their performance on PPT operational tests. English speakers (students who were designated as English only, initially fluent English proficient, or reclassified fluent English proficient) had higher residual means than English learners (effect sizes range from 0.18 to 0.32); students of high SES (i.e., economically advantaged) had higher residual means than low SES students (i.e., economically disadvantaged); these effect sizes range from 0.08 to 0.31. Additionally, students who do not receive special education service had slightly higher residual means than students who receive special education service, with effect sizes ranging from 0.02 to 0.14; and students whose parents or guardians had some postsecondary education had larger residual means than students with parents or guardians with a high school education or lower (effect sizes range from 0.12 to 0.41).

Table 46. Subgroup Residual Summary Statistics: Grade Five Science

	Residual		
Group	**No.**	**Mean**	**SD**
All	8,426	0.00	5.55
Male	4,147	0.14	5.74
Female	4,279	–0.13	5.36
Gender unknown	0	–	–
American Indian	52	0.94	5.18
Asian American	1,049	1.19	5.71
Pacific Islander	41	0.04	5.43
Filipino	259	0.38	6.58
Hispanic	4,426	–0.73	5.23
African American	425	–0.42	5.18
White	2,014	1.01	5.79
Two or more races	160	–0.19	6.04
Ethnicity unknown	0	–	–
No special education services	8,028	0.01	5.54
Special education services	398	–0.12	5.83
Special education services unknown	0	–	–
English only	4,442	0.38	5.63
Initially fluent English proficient	436	0.39	5.97
English learner	2,053	–0.99	5.15
Reclassified fluent English proficient	1,490	0.12	5.56
English proficiency unknown	5	–	–

	Residual		
Group	**No.**	**Mean**	**SD**
Not economically disadvantaged	3,077	0.89	5.95
Economically disadvantaged	5,341	–0.51	5.24
Economic status unknown	8	–	–
Graduate school/Postgraduate training	896	0.89	6.51
College graduate	1,357	0.78	5.68
Some college (includes AA degree)	1,713	0.37	5.38
High school graduate	1,783	–0.49	5.31
Not a high school graduate	1,552	–0.82	4.96
Parent education level unknown	1,125	–0.29	5.71
Metropolitan	8,200	–0.02	5.57
Small town/Rural	226	0.59	4.93

Table 47. Subgroup Residual Summary Statistics: Grade Eight Science

	Residual		
Group	**No.**	**Mean**	**SD**
All	4,863	0.00	6.34
Male	2,418	0.42	6.56
Female	2,444	–0.42	6.08
Gender unknown	1	–	–
American Indian	45	–2.54	6.14
Asian American	568	0.59	6.76
Pacific Islander	21	–1.88	5.68
Filipino	134	1.60	5.67
Hispanic	2,401	–0.96	6.17
African American	246	–1.00	5.85
White	1,348	1.59	6.32
Two or more races	100	0.03	5.22
Ethnicity unknown	0	–	–
No special education services	4,658	0.02	6.31
Special education services	205	–0.56	7.06
Special education services unknown	0	–	–
English only	2,616	0.51	6.30
Initially fluent English proficient	445	0.12	5.59
English learner	630	–1.54	6.47
Reclassified fluent English proficient	1,170	–0.35	6.46
English proficiency unknown	2	-	-
Not economically disadvantaged	1,872	1.20	6.36
Economically disadvantaged	2,988	–0.74	6.21
Economic status unknown	3	–	–
Graduate school/Postgraduate training	536	1.71	6.65
College graduate	886	1.17	6.56
Some college (includes AA degree)	979	–0.06	6.27

	Residual		
Group	**No.**	**Mean**	**SD**
High school graduate	1,082	–0.82	6.23
Not a high school graduate	836	–0.87	6.01
Parent education level unknown	544	–0.52	5.90
Metropolitan	4,585	–0.04	6.36
Small town/Rural	278	0.69	5.91

Table 48. Subgroup Residual Summary Statistics: Biology

	Residual		
Group	**No.**	**Mean**	**SD**
All	4,587	0.00	6.62
Male	2,270	–0.08	6.87
Female	2,317	0.08	6.37
Gender unknown	0	–	–
American Indian	59	–0.84	5.53
Asian American	333	–0.17	6.81
Pacific Islander	32	0.23	6.59
Filipino	128	0.97	6.89
Hispanic	2,358	–0.34	6.34
African American	266	–0.49	6.05
White	1,323	0.62	7.14
Two or more races	88	0.93	6.35
Ethnicity unknown	0	–	–
No special education services	4,396	0.04	6.62
Special education services	191	–0.87	6.58
Special education services unknown	0	–	–
English only	2,579	0.17	6.78
Initially fluent English proficient	307	0.30	6.81
English learner	475	–1.51	5.77
Reclassified fluent English proficient	1,224	0.15	6.47
English proficiency unknown	2	–	–
Not economically disadvantaged	1,910	0.30	7.09
Economically disadvantaged	2,675	–0.21	6.26
Economic status unknown	2	–	–
Graduate school/Postgraduate training	418	0.30	7.63
College graduate	741	0.68	6.80
Some college (includes AA degree)	1,122	0.27	6.68
High school graduate	1,024	–0.52	6.15
Not a high school graduate	797	–0.55	6.43
Parent education level unknown	485	0.08	6.42
Metropolitan	4,297	0.18	6.58
Small town/Rural	290	–2.62	6.69

Student- and School-level Readiness Factors for CBT Implementation

This section of the report addresses the following research questions: What are the major factors that impact students and school readiness for CBT? How do these factors contribute to students' performance on the CBT Tryout?

Survey data collected at the student and at the school level were used to address these research questions. The readiness factors were investigated by means of exploratory and confirmatory factor analyses on the survey data. The impact and merit of these factors were evaluated using factor scores and the residual scores from the previous analysis as described previously in *Differential Impact at the Overall Test Level: Regression and Residual Analyses*.

Data

Student-level Survey

Fifteen questions (questions 1–15) from the student-level survey described earlier were used in this analysis. These questions were designed to collect information on students' technology readiness for CBT and asked students about their exposure to computer-related technologies in general, both inside and outside of school, efficacy in computer-related technologies, reactions to taking a STAR test on computer instead of paper, and previous CBT experiences. This survey is included in Appendix B starting on page 84.

Responses to questions 1–9 and 11–13 were used in the factor analyses. Question 10 was excluded because it measures preference rather than technology readiness. All questions were answered on a 4-point Likert scale except question 13, which is a dichotomous question.

The data used for these analyses include all cases with at least one response to the student-level survey. The data for each test were randomly partitioned into an exploratory subsample and a confirmatory subsample. The total number of responses and the available sample sizes for the exploratory and confirmatory analyses are listed in Table 49.

Table 49. Summary of Factor Analysis Sample Sizes by Science Test

	Total	EFA	CFA
Grade 5	7,743	3,978	3,765
Grade 8	4,924	2,560	2,364
Biology	4,508	2,246	2,262

School-level Survey

Ten questions from the school-level survey (questions 6–15) described earlier were used for the analyses. These questions were designed to be answered by test coordinators or school administrators providing information about the technology environment at the school level, including teachers' professional development in instructional technology and computer assessment, instructional technology use in classrooms, and technology readiness for CBT including general information, hardware, software, and technical support. This survey is also located in Appendix B, starting on page 88.

A total of 102 completed surveys were received. All of the 102 respondents reported themselves as test administrators at the school level, of which 17.6 percent also identified themselves as STAR coordinators for their LEA, 7.8 percent as technology coordinators for their LEA, and 6 percent as

teacher or principal. The remainder did not state their role. Of the 102 respondents, only 52 reported their school identities, representing 5,343 CBT Tryout students from 43 schools in 38 school districts (Table 50). Responses to questions 6–7 and 9–14 were used in the EFA. Responses to questions 8 and 15 were numeric and therefore were not used. All questions were answered on a 4-point Likert scale except question 14, which is a dichotomous question.

Table 50. Number of Districts, Schools, and Students Represented by the 52 School-level Surveys Returned with School Identification

	Districts	Schools	Students
Grade 5	19	22	2,041
Grade 8	10	10	991
Biology	15	15	2,311
Total Tested	38	43	5,343

Methods

Exploratory Factor Analyses

The EFAs were undertaken at the student-level to assess the dimensionality of students' technology readiness for CBT and at the school level to identify dimensions related to the school technology environment. These analyses were conducted using Mplus (Muthen & Muthen, 2009). The loadings were extracted using the maximum likelihood estimation with a promax rotation.

Each solution was evaluated for its ability to produce dimensions that: (a) satisfy Cattell's (1966) scree test; (b) retain two or more items with salient loadings, where loadings > 0.25 are considered salient; (c) yield reasonable internal consistency for items with salient loadings; (d) have eigenvalues greater than 1.0 as specified by the Kaiser-Guttman rule (Brown, 2006); and (e) make psychological sense in terms of parsimonious solutions including mutually exclusive assignment of items to factors, maximum number of items retained, and compatibility with dimensions obtained in other empirical work (Fabrigar, Wegener, MacCallum, & Strahan, 1999).

Confirmatory Factor Analyses

The CFAs were conducted using the student level survey data by fitting data for the reserved subsample (confirmatory subsample) with a four-factor, simple-structure solution, which was the solution obtained from the EFA. Model fit was examined using the CFI, NNFI, and the RMSEA. The CFI ranges from 0 (poor fit) to 1 (good fit). The NNFI is interpreted like the CFI. While there is no single set of evaluation rules agreed upon by all researchers, Hu and Bentler (1999) provide the following guidelines for evaluating fit indices: CFI ≥ 0.95, NNFI ≥ 0.95, and RMSEA < 0.06. No confirmatory analyses were conducted for the school-level survey data sets due to small sample sizes.

Results

Student-level Data

Exploratory Factor Analyses

EFA results suggested the presence of four latent factors. Iterative solutions were tested for one to four factors for the exploratory subsamples of each test and were evaluated against the criteria stated previously. The four-factor solution was found to fit best and to satisfy all of the criteria, in addition to having the most interpretable factors. The same factors were identified across the exploratory subsamples for all three tests.

Using the Biology test as an example, eigenvalues of the first four components were 2.99, 1.42, 1.38, and 1.07 respectively. These four factors accounted for 24.93 percent, 11.83 percent, 11.5

percent, and 8.92 percent of the observed variance respectively, such that the four-factor solution explained 57.17 percent of the total variance.

Salience of loading was found for all 12 questions and no questions were removed from the scale (as recommended by Comrey, 1988). The results were consistent with a simple structure solution where each question loaded significantly (factor loading > 0.25) on only one of the four factors. The final factors, component questions, pattern matrix, and the correlations between factors are presented in Table 51. The correlations between the factors range from 0.19 to 0.37.

The number of questions with salient loadings on each of the four factors and the corresponding interpretation of the factors is as follows:

1. Two questions defined the first factor, which could be called *exposure to computer-related technologies* (e.g., "How often do you send/receive text messages?").
2. Two questions defined the second factor, which could be called *exposure to computer technologies in a nonacademic environment* (e.g., "How often do your other family members use computers at home?").
3. Four questions defined the third factor, which could be called *exposure to computer technologies in academics* (e.g., "How often do you use a computer for all academic study at both school and home?").
4. Four questions defined the fourth factor, which could be called *efficacy and attitude to computer technologies* (e.g., "I can type as quickly and accurately on a computer test as I write on a paper-and-pencil test.").

Table 51. Pattern Matrix for EFA with Promax Rotation of the Student Technology Survey, Biology Test Exploratory Sample (n=2,246)

Question	Question Text	F1	F2	F3	F4
2	How often do you send/receive text messages?	0.95	–0.06	0.01	–0.02
1	How often do you use mobile electronic devices, such as e-reader (including Kindle, Nook, etc.), tablets (including iPad) or smart phones (including iPhone)?	0.59	0.17	–0.02	0.09
7	How often do your other family members use computers at home?	0.05	0.60	0.03	–0.05
6	How often do you use a computer to play games?	–0.21	0.36	0.01	0.25
4	How much time did you spend doing work (including reading, writing and watching a video) for Science class on a computer last year? Include work you do in class and for homework.	–0.05	–0.12	0.67	–0.03
5	How often do you use a computer for all academic study at both school and home?	0.07	0.23	0.49	0.03
3	How often did your science teacher use a computer in the class instruction last year?	–0.05	0.06	0.33	0.00
11	Have you ever taken a school test on a computer before? The test may have been a classroom test, a standardized test, a practice test or quiz, or any other type of test.	0.06	0.07	0.26	0.02
9	I am relaxed when I am working on a computer.	–0.03	0.16	–0.01	0.70
12	I can type as quickly and accurately on a computer test as I write on a paper and pencil test.	0.06	–0.08	0.10	0.59
8	How would you rate your computer skills?	–0.01	0.25	–0.08	0.56
13	Would you rather take tests on a computer or on paper?	0.04	–0.15	0.01	0.53

Note: $r_{F_1,F_2} = .29,\ r_{F_1,F_3} = .19, r_{F_2,F_3} = .34, r_{F_1,F_4} = .25,\ r_{F_2,F_4} = .37, r_{F_3,F_4} = .30$,

F1: Exposure to computer-related technologies

F2: Exposure to computer technologies in nonacademic environment

F3: Exposure to computer technologies in academics

F4: Efficacy and attitude to computer technologies

Confirmatory Factor Analyses

The CFAs for the responses to the 12 questions supported the four-factor simple-structure model, as shown in Table 52. The fit statistics (see Table 53) suggested adequate model-data fit for the confirmatory subsample of each test.

Table 52. Factor Loadings for Biology CFA 4-factor Simple-Structure Model for Responses on Nine Questions from the Confirmatory Sample (n=2,262)

F1: Exposure to computer-related technologies		
1	How often do you use mobile electronic devices, such as e-reader (including Kindle, Nook, etc.), tablets (including iPad) or smart phones (including iPhone)?	0.78
2	How often do you send/receive text messages?	0.73
F2: Exposure to computer technologies in non-academic environment		
6	How often do you use a computer to play games?	0.48
7	How often do your other family members use computers at home?	0.52
F3: Exposure to computer technologies in academics		
5	How often do you use a computer for all academic study at both school and home?	0.75
4	How much time did you spend doing work (including reading, writing and watching a video) for Science class on a computer last year? Include work you do in class and for homework.	0.37
3	How often did your science teacher use a computer in the class instruction last year?	0.30
11	Have you ever taken a school test on a computer before? The test may have been a classroom test, a standardized test, a practice test or quiz, or any other type of test.	0.29
F2: Efficacy and attitudes to computer technologies		
9	I am relaxed when I am working on a computer.	0.73
12	I can type as quickly and accurately on a computer test as I write on a paper and pencil test.	0.65
8	How would you rate your computer skills?	0.62
13	Would you rather take tests on a computer or on paper?	0.47

Table 53. CFA Model Fit Statistics

	CFI	NNFI	RMSEA
Grade 5	0.94	0.93	0.04
Grade 8	0.91	0.91	0.05
Biology	0.94	0.94	0.05

School-level Data

The results of the EFA suggested that two factors might be extracted. Eigenvalues of the first two components were 3.15 and 1.50 respectively. These two factors accounted for 45 percent and 21.43 percent of the observed variance respectively, such that the two-factor solution explained 66.43 percent of the total variance. Salience of loading was found for seven of the eight questions and the nonsalient question was removed from the scale (as recommended by Comrey, 1988). The remaining seven

questions (questions 6, 7, 9, 10, 12, 13, and 14) represent a simple structure with the two factors. The final factors, component questions, pattern matrix, and the correlations between factors scores are presented in Table 54. Based on the loadings, results indicate the following:

1. Four questions defined the first factor, which could be called *teachers' experience with technology* (e.g., "What percentage of Science teachers at your school use computers at least once a week as part of their classroom instruction?").
2. Three questions defined the second factor, which could be called *school computer testing experience* (e.g., "How prepared is your school for administering computer-based tests?").

The correlation between the two factors was 0.44.

Table 54. Pattern Matrix of the EFA from the School-level Survey (n=102)

Question	Question Text	F1	F2
9	What percentage of Science teachers at your school use computers at least once a week as part of their classroom instruction?	0.89	–0.23
10	Estimate the percentage of Science teachers at your school who assign homework at least once a week that requires the use of computers (e.g., research, simulation, multimedia presentation, and online collaboration)?	0.85	–0.03
12	During the past school year, how often did teachers at your school use instructional technology with students for activities such as research, multimedia, simulations, data interpretation, communications, and collaboration?	0.54	0.23
6	During the past school year, how many hours of instructional technology professional development did teachers at your school complete, on average?	0.52	0.19
14	Did your school have experience administering computer-based tests before this tryout?	–0.13	0.91
13	How prepared is your school for administering computer-based tests? (including infrastructure, computer assessment software, administrators', teachers' and students' technology readiness)	–0.07	0.75
7	Estimate what percentage of teachers in your school had experience with computer-based assessment for the subjects they teach.	0.18	0.51

Note: $r_{F_1,F_2} = .44$

F1: Teachers' experience with technology

F2: School computer testing experience

Student Performance by Levels of Technology Readiness

In order to investigate how the factors appear to contribute to students' performance on the CBT Tryout, factor scores based on the results of the confirmatory factor analysis model were estimated for each student given their responses on the survey questions. The factor score is a numerical value estimating a person's standing on a latent factor; for example, a higher factor score on "Computer Efficacy" indicates higher computer efficacy.

For use in further analysis with residuals, the scores for each factor were scaled to have a mean of 0 and standard deviation of 1. Table 55, Table 57, and Table 59 present the summary statistics of the scaled factor scores for the total group and each demographic subgroup. Using these scaled factor scores, each factor is dichotomized at zero and students are grouped into two levels on each factor, such as "High Computer Efficacy" versus "Low Computer Efficacy." Residual scores obtained in the residual analysis as described in *Differential Impact at the Overall Test Level: Regression and*

Residual Analyses are averaged by factor level, and residual means for students in the high and low groups are compared to investigate the effects of technology factors on students' CBT performance.

Table 56, Table 58, and Table 60 present the summary statistics of the residual scores by level on each factor and the results of independent samples t-test of the mean residual scores between high and low levels. The effect sizes (Cohen, 1988) are also provided. Figure E.2 through Figure E.13, starting on page 140 in Appendix E, present the cumulative distributions of residual scores for high and low levels on each factor for each test. Discrepancy between the cumulative distributions for the high and low levels of each factor indicates the extent to which that factor differentiates students' performance on the CBT Tryout.

The data in Table 55, Table 57, and Table 59 show that the summary statistics for the factor scores seem to follow a pattern across three tests. For example, on all four factors, English-only students had higher means than English learners; students with high SES (not economically disadvantaged) had higher means than students with low SES (economically disadvantaged); students who do not receive special education services had higher means than students who receive special education services; and students whose parents or guardians have higher education levels had larger means than students with parents or guardians with high school or lower education levels. These results indicate that, among the students participating in the CBT Tryout, English speakers, high SES students, students who do not receive special education services, and students with more highly educated parents or guardians report greater exposure to and higher efficacy in using the computer technologies.

As noted previously, to investigate the possible influence of these factors on students' performance on the CBT Tryout, the residual scores on the CBT Tryout were compared by level of each factor, that is, between the group with high factor scores and the group with low factor scores, as shown in Table 56, Table 58, and Table 60, and in Figure E.2 through Figure E.13.

Results indicate some between-level differences in residual scores, although the differences were small and varied somewhat by test. Not surprisingly, the most salient factor for all three tests was exposure to computers in academic settings; that is, the group that reported greater exposure to computers in academics (F3) had higher residual scores than those with lower levels of exposure. The groups with more exposure to computers in nonacademic environments (F2) or with high efficacy and attitude toward computer technologies (F4) also tended to have higher residual scores, but the magnitude of difference varied across the three tests. Exposure to computer-related technologies was inconsistent; the group with low exposure to computer-related technologies, F1, tended to have a higher mean in residual scores for grade five science and Biology but not for grade eight science.

Taken together, the results imply that students tend to perform better on computer-based tests when they have more exposure to computers in either academic or nonacademic environments or when they have higher efficacy and better attitudes toward computer technologies. However, exposure to computer-related technologies, such as tablets or smart phones, may not be directly helpful in improving computer-based testing performance.

Table 55. Grade Five Science Factor Statistics by Subgroup

		F1		F2		F3		F4	
	No.	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Total	7,519	0.00	1.00	0.00	1.00	0.00	1.00	0.00	1.00
Male	3,597	–0.07	1.01	–0.03	1.02	0.02	1.01	0.03	1.03
Female	3,719	0.06	0.98	0.02	0.98	–0.03	0.98	–0.04	0.97
Gender unknown	203	0.13	1.01	0.15	1.01	0.16	1.01	0.17	1.03

		F1		F2		F3		F4	
	No.	Mean	SD	Mean	SD	Mean	SD	Mean	SD
American Indian	44	–0.02	1.22	0.04	1.11	–0.02	1.22	–0.15	1.15
Asian American	904	–0.02	0.95	0.29	0.95	0.22	0.94	0.15	1.02
Pacific Islander	31	–0.21	1.04	0.03	0.95	0.00	1.07	–0.10	1.03
Filipino	210	0.08	0.91	0.17	0.90	0.23	0.89	0.24	0.94
Hispanic	3,917	–0.09	1.02	–0.13	1.02	–0.16	1.02	–0.12	0.99
African American	344	0.18	0.95	0.05	0.98	0.14	0.95	0.17	0.98
White	1,726	0.15	0.97	0.08	0.95	0.18	0.93	0.12	0.98
Two or more races	140	0.10	0.92	0.05	0.99	0.08	0.94	0.00	0.96
Ethnicity unknown	203	0.13	1.01	0.15	1.01	0.16	1.01	0.17	1.03
No special education services	6,752	0.01	1.00	0.00	1.00	0.00	1.00	0.01	1.00
Special education services	564	–0.11	0.95	–0.09	1.03	–0.12	0.96	–0.17	1.00
Special education services unknown	203	0.13	1.01	0.15	1.01	0.16	1.01	0.17	1.03
English only	3,867	0.09	0.99	0.01	0.98	0.08	0.97	0.05	0.99
Initially fluent English proficient	385	0.03	0.96	0.18	0.95	0.08	0.99	0.11	1.02
English learner	1,784	–0.16	0.99	–0.11	1.04	–0.21	1.02	–0.19	0.99
Reclassified fluent English proficient	1,276	–0.08	1.01	0.03	0.98	0.00	1.02	0.03	1.00
English proficiency unknown	207	0.15	1.01	0.17	1.02	0.19	1.02	0.20	1.03
Not economically disadvantaged	2,657	0.13	0.94	0.14	0.95	0.17	0.93	0.10	0.99
Economically disadvantaged	4,653	–0.08	1.02	–0.09	1.02	–0.10	1.02	–0.07	1.00
Economic status unknown	209	0.11	1.02	0.13	1.01	0.15	1.01	0.16	1.03
Graduate school/Postgraduate training	760	0.16	0.90	0.25	0.89	0.26	0.89	0.18	0.96
College graduate	1,177	0.10	0.97	0.16	0.96	0.18	0.94	0.12	0.99
Some college (includes AA degree)	1,497	0.05	1.01	–0.03	0.98	0.04	1.00	0.07	1.01
High school graduate	1,564	0.00	0.99	–0.07	0.99	–0.05	0.96	–0.03	0.96
Not a high school graduate	1,358	–0.20	1.04	–0.17	1.05	–0.25	1.06	–0.19	1.02
Parent education level unknown	1,163	–0.04	1.00	0.01	1.03	–0.05	1.02	–0.06	1.01
Metropolitan	7,283	0.00	1.00	0.00	1.00	0.00	1.00	0.00	1.00
Small town/Rural	236	0.02	1.03	–0.04	1.02	0.02	0.91	–0.05	0.90

F1: Exposure to computer-related technologies

F2: Exposure to computer technologies in nonacademic environment

F3: Exposure to computer technologies in academics

F4: Efficacy and attitude to computer technologies

Table 56. Grade Five Science Regression Residuals Grouped by Factors 1–4

	Residual					Group Comparisons		
	No.	Mean	SD	Min	Max	t-value	*p*-value	Cohen's *d*
Total	7,058	0.28	5.31	–30.84	23.99	–	–	–
F1 low	3,430	0.43	5.13	–18.97	23.99	2.21	0.027	0.05
F1 high	3,628	0.15	5.48	–30.84	23.92			
F2 low	3,376	0.27	5.15	–28.81	23.92	–0.14	0.889	0.00
F2 high	3,682	0.29	5.46	–30.84	23.99			
F3 low	3,489	0.10	5.16	–28.81	23.92	–2.91	0.004	–0.07
F3 high	3,569	0.46	5.45	–30.84	23.99			
F4 low	3,502	0.20	5.16	–19.32	23.92	–1.32	0.189	–0.03
F4 high	3,556	0.36	5.45	–30.84	23.99			

F1: Exposure to computer-related technologies

F2: Exposure to computer technologies in nonacademic environment

F3: Exposure to computer technologies in academics

F4: Efficacy and attitude to computer technologies

Table 57. Grade Eight Science Factor Statistics by Subgroup

		F1		F2		F3		F4	
	No.	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Total	4,816	0.00	1.00	0.00	1.00	0.00	1.00	0.00	1.00
Male	2,165	–0.11	1.04	0.00	1.03	0.06	1.03	0.06	1.03
Female	2,202	0.11	0.95	0.01	0.95	–0.05	0.96	–0.06	0.96
Gender unknown	449	–0.01	1.02	–0.07	1.08	–0.04	1.05	–0.03	1.04
American Indian	31	–0.27	1.30	–0.26	1.24	–0.40	1.15	–0.03	0.98
Asian American	488	–0.13	1.07	0.34	0.92	0.37	0.97	0.19	1.05
Pacific Islander	21	–0.12	1.25	0.29	1.05	0.30	1.06	0.28	1.09
Filipino	121	0.24	0.81	0.36	0.82	0.34	0.91	0.17	1.05
Hispanic	2,180	–0.10	1.00	–0.21	0.98	–0.27	0.97	–0.19	0.96
African American	220	0.02	0.94	0.02	1.00	0.12	0.91	0.16	0.92
White	1,221	0.21	0.93	0.21	0.95	0.28	0.92	0.20	0.99
Two or more races	86	–0.03	1.11	0.10	1.12	0.26	1.01	0.25	0.96
Ethnicity unknown	448	–0.01	1.02	–0.07	1.08	–0.03	1.05	–0.02	1.04
No special education services	4,047	0.03	0.99	0.03	0.98	0.03	0.99	0.03	0.99
Special education services	321	–0.32	1.04	–0.29	1.05	–0.30	1.01	–0.32	1.04
Special education services unknown	448	–0.01	1.02	–0.07	1.08	–0.03	1.05	–0.02	1.04
English only	2,365	0.09	0.98	0.06	1.01	0.12	0.99	0.09	1.00
Initially fluent English proficient	399	0.01	0.97	0.04	0.93	0.00	0.95	–0.03	0.95
English learner	569	–0.33	1.05	–0.26	1.05	–0.37	1.03	–0.29	1.00
Reclassified fluent English proficient	1,034	–0.03	0.98	0.03	0.91	–0.05	0.95	–0.03	0.96
English proficiency unknown	449	–0.01	1.02	–0.07	1.08	–0.03	1.05	–0.02	1.04
Not economically disadvantaged	1,706	0.19	0.93	0.19	0.98	0.23	0.98	0.15	1.02
Economically disadvantaged	2,661	–0.12	1.02	–0.11	0.98	–0.14	0.98	–0.09	0.97
Economic status unknown	449	–0.01	1.02	–0.07	1.08	–0.03	1.05	–0.02	1.04
Graduate school/Postgraduate training	486	0.17	0.99	0.34	0.99	0.37	0.97	0.21	1.03
College graduate	808	0.22	0.93	0.29	0.95	0.32	0.95	0.25	1.00
Some college (includes AA degree)	883	0.05	0.95	–0.02	0.95	0.05	0.93	0.04	0.95
High school graduate	961	–0.12	1.05	–0.12	0.97	–0.15	0.99	–0.11	1.00
Not a high school graduate	755	–0.14	0.98	–0.19	0.96	–0.27	0.95	–0.19	0.93
Parent education level unknown	923	–0.08	1.02	–0.13	1.06	–0.14	1.04	–0.09	1.02
Metropolitan	4,551	0.00	1.00	0.00	1.00	0.01	1.00	0.00	1.00
Small town/Rural	265	–0.05	0.97	–0.05	0.98	–0.14	0.92	–0.03	0.96

F1: Exposure to computer-related technologies

F2: Exposure to computer technologies in non-academic environment

F3: Exposure to computer technologies in academics

F4: Efficacy and attitude to computer technologies

Table 58. Grade Eight Science Regression Residuals Grouped by Factors 1–4

	Residual					Group Comparisons		
	No.	Mean	SD	Min	Max	t-value	*p*-value	Cohen's *d*
Total	4,217	0.22	6.14	–28.78	28.55	–	–	–
F1 low	1,505	–0.01	6.23	–27.02	19.15	–1.84	0.066	–0.06
F1 high	2,712	0.35	6.08	–28.78	28.55			
F2 low	2,040	–0.36	6.10	–27.02	21.35	–5.93	< 0.001	–0.18
F2 high	2,177	0.76	6.12	–28.78	28.55			
F3 low	2,104	–0.53	6.17	–27.02	22.83	–7.99	< 0.001	–0.25
F3 high	2,113	0.97	6.01	–28.78	28.55			
F4 low	2,088	–0.27	6.23	–28.78	22.83	–5.16	< 0.001	–0.16
F4 high	2,129	0.70	6.01	–24.78	28.55			

F1: Exposure to computer-related technologies

F2: Exposure to computer technologies in nonacademic environment

F3: Exposure to computer technologies in academics

F4: Efficacy and attitude to computer technologies

Table 59. Biology Factor Statistics by Subgroup

		F1		F2		F3		F4	
	No.	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Total	4,428	0.00	1.00	0.00	1.00	0.00	1.00	0.00	1.00
Male	2,135	–0.08	1.07	–0.05	1.05	0.04	1.04	0.04	1.05
Female	2,189	0.08	0.93	0.05	0.95	–0.04	0.95	–0.04	0.94
Gender unknown	104	–0.03	0.93	0.04	1.01	0.07	1.07	–0.04	1.13
American Indian	58	0.01	0.97	0.03	1.01	0.01	0.94	–0.03	0.93
Asian American	304	–0.19	1.17	0.21	1.01	0.29	1.01	0.20	1.04
Pacific Islander	33	–0.18	1.10	–0.11	0.86	–0.08	0.75	–0.02	0.84
Filipino	110	0.01	0.98	0.31	0.97	0.37	0.94	0.30	0.94
Hispanic	2,189	–0.08	0.99	–0.16	0.98	–0.23	0.97	–0.17	0.97
African American	237	0.15	0.91	0.02	0.96	0.08	0.96	0.11	0.99
White	1,304	0.14	0.98	0.16	1.01	0.25	0.98	0.17	1.01
Two or more races	89	0.25	0.82	0.31	0.99	0.34	0.92	0.33	0.92
Ethnicity unknown	104	–0.03	0.93	0.04	1.01	0.07	1.07	–0.04	1.13
No special education services	4,129	0.02	0.99	0.00	0.99	0.01	0.99	0.01	0.99
Special education services	193	–0.32	1.14	–0.11	1.13	–0.16	1.16	–0.23	1.12
Special education services unknown	106	–0.01	0.93	0.05	1.01	0.07	1.06	–0.03	1.12
English only	2,493	0.10	0.98	0.08	1.01	0.14	0.99	0.12	1.00
Initially fluent English proficient	289	–0.05	1.01	–0.09	0.92	–0.09	0.91	–0.07	0.95
English learner	441	–0.32	1.03	–0.27	1.01	–0.43	0.97	–0.37	0.94
Reclassified fluent English proficient	1,098	–0.09	1.00	–0.06	0.98	–0.13	0.98	–0.10	0.98
English proficiency unknown	107	–0.01	0.93	0.05	1.00	0.06	1.06	–0.03	1.12
Not economically disadvantaged	1,838	0.14	0.96	0.14	0.99	0.20	0.98	0.15	1.01
Economically disadvantaged	2,483	–0.10	1.02	–0.11	0.99	–0.15	0.98	–0.11	0.97
Economic status unknown	107	–0.01	0.93	0.05	1.00	0.06	1.06	–0.03	1.12
Graduate school/Postgraduate training	395	0.08	1.04	0.28	1.02	0.34	0.97	0.21	1.04
College graduate	710	0.10	0.96	0.15	0.98	0.22	0.97	0.11	0.99
Some college (includes AA degree)	1,078	0.12	0.98	0.09	1.00	0.13	0.99	0.13	1.00

		F1		F2		F3		F4	
	No.	Mean	SD	Mean	SD	Mean	SD	Mean	SD
High school graduate	945	0.00	0.99	–0.10	0.98	–0.13	0.97	–0.05	0.97
Not a high school graduate	733	–0.20	1.01	–0.23	0.97	–0.34	0.96	–0.27	0.96
Parent education level unknown	567	–0.14	1.02	–0.10	1.00	–0.10	1.00	–0.08	0.99
Metropolitan	4,064	–0.01	1.00	–0.05	0.98	–0.02	0.99	0.00	1.00
Small town/Rural	364	0.15	0.97	0.61	1.04	0.18	1.04	0.01	0.99

F1: Exposure to computer-related technologies

F2: Exposure to computer technologies in nonacademic environment

F3: Exposure to computer technologies in academics

F4: Efficacy and attitude to computer technologies

Table 60. Biology Regression Residuals Grouped by Factors 1–4

	Residual					Group Comparisons		
	No.	Mean	SD	Min	Max	t-value	*p*-value	Cohen's *d*
Total	4,020	0.26	6.50	–28.01	26.78	–	–	–
F1 low	1,497	0.56	6.65	–28.01	21.40	2.24	0.025	0.07
F1 high	2,523	0.09	6.41	–23.08	26.78			
F2 low	2,020	–0.01	6.47	–27.68	24.06	–2.67	0.008	–0.08
F2 high	2,000	0.54	6.53	–28.01	26.78			
F3 low	2,014	–0.19	6.34	–27.68	21.40	–4.42	< 0.001	–0.14
F3 high	2,006	0.72	6.63	–28.01	26.78			
F4 low	2,056	–0.01	6.32	–27.68	21.40	–2.75	0.006	–0.09
F4 high	1,964	0.55	6.68	–28.01	26.78			

F1: Exposure to computer-related technologies

F2: Exposure to computer technologies in nonacademic environment

F3: Exposure to computer technologies in academics

F4: Efficacy and attitude to computer technologies

The potential impact of school-level technology readiness on students' performance on the CBT Tryout was examined for Biology only. Most of the grade six students who took the grade five science test and the grade nine students who took grade eight science test had just changed schools when they participated in the CBT Tryout; consequently, the school-level technology readiness at their current school would have contributed little to their performance.

For the Biology school-level technology readiness analysis, factor scores were estimated for each school represented by respondents to the school-level survey. As with the student-level analyses, scores for each factor were scaled to have a mean of 0 and a standard deviation of 1. For the seven schools that were represented by more than one respondent, the respondents' scaled factor scores were averaged to provide a single score for each school. Using these scaled factor scores, each factor was dichotomized at zero and schools were grouped into two levels on each factor. Residual scores obtained in the residual analysis as described in *Differential Impact at the Overall Test Level: Regression and Residual Analyses* were then averaged by factor level and residual means for students in the high and low school groups were compared to investigate the effects of school-level technology factors on students' CBT performance.

Table 61 provides the summary statistics of the residual scores by level on each factor and the results of independent sample t-tests of the mean residual scores between high and low levels. The results indicated that students in schools where teachers had a high level of experience with technology

(F1) had higher residual scores than those in schools with a lower level of teachers' experience with technology, but the magnitude of that difference was small; students in schools with more CBT experience (F2) had slightly lower residual scores than those in schools with less CBT experience.

Table 61. Biology Regression Residuals Grouped by Factors 1 and 2 on School-level Technology Readiness

		Residual				Group Comparisons		
	No.	**Mean**	**SD**	**Min**	**Max**	**t-value**	***p*-value**	**Cohen's *d***
All	4,587	0.00	6.62	–31.01	26.78	–	–	–
F1 Low	651	0.05	6.93	–29.89	26.78	–2.37	0.018	–0.12
F1 High	1,044	0.84	6.49	–31.01	24.06			
F2 Low	1,124	0.47	6.70	–31.01	24.06	0.67	0.500	0.03
F2 High	886	0.27	6.41	–25.22	26.78			

F1: Teachers' experience with technology

F2: School computer testing experience

The results imply that students tend to perform better than expected on computer-based tests when they are in a school where teachers use computers more often in instruction and have more professional development in instructional technology. However, students in a school with more computer testing experience (F2) do not necessarily perform better than students in a school with less computer testing experience. Figure E.14 and Figure E.15 on page 142 in Appendix E, which present the cumulative distributions of residual scores of the two levels on each school-level factor for the Biology test, reflect the results provided in Table 60. Specifically, Figure E.14 shows some discrepancy between cumulative distribution curves of high and low factor score groups on teachers who have experience with technology (F1) and Figure E.15 shows very little discrepancy between high and low factor score groups on CBT experience (F2).

It should be noted that these results are based on a relatively small subset of schools that participated in the CBT for Biology. The response rate on the school-level survey was low and of those who did respond, half did not report their school identification numbers and therefore could not be associated with the students in the schools. Of the 56 schools that administered the Biology CBT, only 15 schools were represented by respondents to the school-level technology survey. Therefore, the generalizability of these results is very limited and should be interpreted with caution.

Conclusions of the Psychometric Studies

The psychometric studies were conducted to investigate four research questions:

1. What are the statistical properties of the CBT Tryout forms and items comprising the forms?
2. Do TE items measure the same underlying construct as traditional MC items for each of the three grades?
3. Does CBT lead to differential impact on student performance at the item and test level for subgroups of interest classified by key demographic variables?
4. What are the major factors that impact students and school readiness for CBT? How do these factors contribute to students' performance on the CBT Tryout?

The first research question investigated the statistical properties of the CBT Tryout forms and the items which comprise the forms. The results of the classical item- and test-level analyses indicated that the quality of the CBT forms was generally supported by desirable psychometric properties at both the test and the item levels. For example, the CBT forms were shown to be reliable with high coefficient of internal consistency; the difficulty and discrimination of the CBT forms were comparable to the corresponding PPT forms as item difficulties and item discriminations were highly correlated between the two testing modes. Some differences were observed in omit rates between CBT and PPT administrations of the MC items, and between TE and MC items, as well as between English learners and non–English learners. However, these omit rates were generally less than the conventionally accepted omit rate of five percent.

The second research question examined the factor structure or dimensionality of the CBT Tryout assessments. The results of factor analyses supported a unidimensional interpretation of the CBT Tryout forms that include both TE and traditional MC items.

The third question addressed whether the performance of students from different demographic subgroups was differentially affected by CBT administration compared to PPT at the item and test levels. At the item level, DIF analyses showed that some of the MC items, not flagged as C-DIF in PPT, were flagged as C-DIF when administered in the CBT Tryout. This finding suggests that CBT administration may differentially impact students' performance depending on gender, ethnicity, English fluency, disability status, and area of residency. However, since DIF analyses for the operational PPT do not include comparisons for area of residency, it is unknown if these items would have been flagged as C-DIF for these subgroups on the PPT.

At the test level, the comparison of residual score distributions between subgroups showed small differences in student performance on each test with similar patterns of differences between subgroups across tests. That is, English speakers performed better than expected on the CBT Tryout in contrast to the English learners when controlling for their performance on PPT operational tests. Additionally, students who did not receive special education services performed better on the CBT Tryout than students who receive special education services, and students with high SES performed better on the CBT Tryout than low SES students. However these findings should be interpreted with caution due to the limitations of the study, including considerations of motivation (no-stakes), timing (i.e., students were administered the CBT Tryout forms several months after they completed the course in the subject), and the design of the study, which did not include the counterbalanced administration of paper versions of the items and forms.

The fourth question focused on the major factors that impact student and school readiness for CBT and how these factors contributed to students' performance on the CBT Tryout. The student-level survey data supported a four-factor simple-structure model with factors one to four interpreted as

exposure to computer-related technologies, exposure to computer technologies in a nonacademic environment, exposure to computer technologies in academics, and efficacy and attitude to computer technologies, respectively.

Comparison of residuals for students classified as low or high on each factor revealed some differences between the groups; for example, students with high exposure to computer technologies in academic or nonacademic environments tended to perform better than expected on the CBT Tryout given their spring CST scores. However, the effect sizes of these differences are relatively small, suggesting that, individually, these factors may not have contributed in a meaningful way to students' performance on the CBT Tryout.

For the school-level survey, two factors are identified: teachers' experience with technology and school computer testing experience. Comparison of Biology CBT residuals for schools classified as low or high on each factor suggested that students in schools where teachers have higher levels of experience with technology tend to perform better on computer-based tests. However, the observed difference was relatively small and the results should be interpreted with caution given the limited sample that resulted from low response rates and missing school identification numbers in school-level survey responses. The CBT residual scores were not compared between schools with low and high factor scores for grade five and grade eight science tests because of the recent school change for students taking the tests.

Appendixes

Appendix A—Participation

Schools That Tested Students as Part of the CBT Tryout

Schools That Tested Students				
CD code	**District Name**	**CDS code**	**School Name**	**Charter/ Ind. Testing Charter?**
1964212	ABC Unified	19642126071369	Bragg Elementary	
1964212	ABC Unified	19642121930056	Cerritos High	
0161119	Alameda City Unified	01611190130229	Alameda High	
0161119	Alameda City Unified	01611196090054	Lincoln Middle	
1975713	Alhambra Unified	19757131930163	Alhambra High	
1975713	Alhambra Unified	19757136011035	Garfield Elementary	
3066423	Anaheim City	30664236027387	Henry (Patrick) Elementary	
3066423	Anaheim City	30664230113712	Orange Grove Elementary	
3066431	Anaheim Union High	30664313038239	Western High	
1964246	Antelope Valley Union High	19642461995844	Lancaster High	
3675077	Apple Valley Unified	36750776107346	Sandia Elementary	
1964287	Baldwin Park Unified	19642876110043	Santa Fe Elementary	
1964287	Baldwin Park Unified	19642871938166	Sierra Vista High	
3667637	Bear Valley Unified	36676376105936	Big Bear Elementary	
1964303	Bellflower Unified	19643036011704	Foster (Stephen) Elementary	
1964303	Bellflower Unified	19643036011696	Ramona Elementary	
4870524	Benicia Unified	48705244831004	Benicia High	
4870524	Benicia Unified	48705246050983	Benicia Middle	
4970623	Bennett Valley Union Elementary	49706236098248	Strawberry Elementary	
1964311	Beverly Hills Unified	19643116011779	Horace Mann Elementary	
1964329	Bonita Unified	19643291937739	San Dimas High	
2765979	Bradley Union Elementary	27659796026017	Bradley Elementary	
1363107	Calipatria Unified	13631076008395	Young Middle	
3066464	Capistrano Unified	30664643030574	Aliso Niguel High	
3066464	Capistrano Unified	30664640113381	San Juan Hills High	
1964352	Centinela Valley Union High	19643521935048	Leuzinger High	
3066472	Centralia Elementary	30664726027684	Danbrook Elementary	
2075606	Chawanakee Unified	20756060117010	Minarets High	
2075606	Chawanakee Unified	20756066024079	North Fork Elementary	
2075606	Chawanakee Unified	20756066024103	Spring Valley Elementary	
3667686	Colton Joint Unified	36676866035604	Crestmore Elementary	
3667686	Colton Joint Unified	36676863630399	Washington High	
1964444	Culver City Unified	19644446057608	Culver City Middle	
4275010	Cuyama Joint Unified	42750104231205	Cuyama Valley High	
5772678	Davis Joint Unified	57726780119578	Da Vinci Charter Academy	Charter
3768056	Del Mar Union Elementary	37680566117923	Sage Canyon	
3768056	Del Mar Union Elementary	37680566120596	Torrey Hills	
0861820	Del Norte County Unified	08618200833004	Del Norte High	
3367058	Desert Sands Unified	33670580118885	Shadow Hills High	
5171373	East Nicolaus Joint Union High	51713735132758	East Nicolaus High	
0961853	El Dorado Union High	09618530930214	Shenandoah High	Charter
1964519	El Monte Union High	19645191932664	El Monte High	
3467314	Elk Grove Unified	34673143432572	Elk Grove High	
3768098	Escondido Union	37680986038210	Mission Mid	
3667702	Etiwanda Elementary	36677020102947	Day Creek Intermediate	

Schools That Tested Students				
CD code	**District Name**	**CDS code**	**School Name**	**Charter/ Ind. Testing Charter?**
3667702	Etiwanda Elementary	36677026119630	Heritage Intermediate	
3166829	Eureka Union	31668296111702	Ridgeview Elementary	
4369435	Evergreen Elementary	43694356067193	Cedar Grove Elementary	
4369435	Evergreen Elementary	43694356047161	Whaley (O. B.) Elementary	
4569989	Fall River Joint Unified	45699894533600	Fall River Junior-Senior High	
3467330	Folsom-Cordova Unified	34673303431533	Cordova High	
3667710	Fontana Unified	36677100120758	Jurupa Hills High	
3667710	Fontana Unified	36677106106470	Locust Elementary	
1010108	Fresno County Office of Education	10101081030337	Fresno County Court	
1062166	Fresno Unified	10621666006118	Burroughs Elementary	
1062166	Fresno Unified	10621666006431	Powers-Ginsburg Elementary	
1062166	Fresno Unified	10621661030642	School of Unlimited Learn	Charter
1964550	Garvey Elementary	19645506013544	Willard (Frances E.) Elementary	
4369484	Gilroy Unified	43694846098214	South Valley Middle	
1964568	Glendale Unified	19645681996131	Clark (A.W.) Magnet High	
1964568	Glendale Unified	19645681995497	Verdugo Academy	
5071084	Gratton Elementary	50710840120089	Gratton Charter	Charter
4770326	Grenada Elementary	47703266050777	Grenada Elementary	
3768130	Grossmont Union High	37681303731809	El Capitan High	
3768130	Grossmont Union High	37681303734548	Monte Vista High	
1973445	Hacienda la Puente Unified	19734451938935	Valley Alter. High, Cont.	
1663925	Hanford Joint Union High	16639251630169	Hanford West High	
3367082	Hemet Unified	33670826112007	Dartmouth Middle	
1964600	Hermosa Beach City Elementary	19646006095434	Hermosa Valley Elementary	
5672462	Hueneme Elementary	56724626055073	Hueneme Elementary	
5672462	Hueneme Elementary	56724626055107	Richard Bard Elementary	
1363164	Imperial Unified	13631641331115	Imperial Ave Holbrook High	
3073650	Irvine Unified	30736506089445	Greentree Elementary	
3073650	Irvine Unified	30736506106850	Meadow Park	
3073650	Irvine Unified	30736506100861	Northwood Elementary	
3367090	Jurupa Unified	33670903337136	Rubidoux High	
3367090	Jurupa Unified	33670906032239	Van Buren Elementary	
1073999	Kerman Unified	10739996119978	Liberty Elementary	
1563529	Kern Union High	15635291532605	Foothill High	
1563529	Kern Union High	15635291533330	Highland High	
1663941	Kings River-Hardwick Union Elementary	16639416010474	Kings River-Hardwick Elementary	
5071142	Knights Ferry Elementary	50711426052609	Knights Ferry Elementary	
1563560	Lamont Elementary	15635606009674	Alicante Avenue Elementary	
1864139	Lassen Union High	18641390106385	Diamond Mountain Charter High	Charter
2465730	Le Grand Union High	24657302433001	Le Grand High	
2375218	Leggett Valley Unified	23752182332724	Leggett Valley High	
1964725	Long Beach Unified	19647251995539	Calif Acad Math and Science	
3166845	Loomis Union Elementary	31668456031132	Loomis Elementary	
1910199	Los Angeles County Office of Education	19101990109942	Los Angeles International Cha	IT Charter
1964733	Los Angeles Unified	19647331931518	George S. Patton Continuation	
1964733	Los Angeles Unified	19647331935352	Los Angeles Senior High	
1964733	Los Angeles Unified	19647336061550	Olive Vista Middle	
1964733	Los Angeles Unified	19647331930551	San Antonio Continuation	
1964733	Los Angeles Unified	19647336019228	Sierra Vista Elementary	

Schools That Tested Students				
CD code	**District Name**	**CDS code**	**School Name**	**Charter/ Ind. Testing Charter?**
1964733	Los Angeles Unified	19647330122366	Social Justice Leadership Acad	
1964774	Lynwood Unified	19647741935428	Vista High (Continuation)	
3968593	Manteca Unified	39685933932001	East Union High	
5872736	Marysville Joint Unified	58727365830013	Lindhurst High	
5872736	Marysville Joint Unified	58727366056741	Olivehurst Elementary	
4373387	Milpitas Unified	43733876047641	Pomeroy (Marshall) Elementary	
4373387	Milpitas Unified	43733876047591	Spangler (Anthony) Elementary	
5071167	Modesto City Elementary	50711676052690	Franklin Elementary	
5071167	Modesto City Elementary	50711676105670	Kirschen (Harriette) Elem	
5071167	Modesto City Elementary	50711676093512	Sonoma Elementary	
1563677	Mojave Unified	15636770113837	Hacienda Elementary	
1062323	Monroe Elementary	10623236006993	Monroe Elementary	
1964790	Monrovia Unified	19647901996404	Mountain Park	
3367124	Moreno Valley Unified	33671243331071	Vista del Lago High	
3667777	Morongo Unified	36677773636743	Twentynine Palms High	
1563685	Muroc Joint Unified	15636851530997	Boron Junior-Senior High	
1563685	Muroc Joint Unified	15636851531987	Desert Junior-Senior High	
1563685	Muroc Joint Unified	15636856009906	West Boron Elementary	
3768221	National Elementary	37682216108559	Rancho de la Nacion	
3475283	Natomas Unified	34752830102277	Witter Ranch Elementary	
2966357	Nevada Joint Union High	29663572930030	Sierra Mountain High	
2966357	Nevada Joint Union High	29663570112367	William & Marian Ghidotti High	
3968627	New Jerusalem Elementary	39686276119309	Delta Charter	IT Charter
1964840	Norwalk-La Mirada Unified	19648406020853	Chavez (Cesar) Elementary	
3667819	Ontario-Montclair Elementary	36678196036271	Howard Elementary	
3010306	Orange County Department of Education	30103063030632	OCCS:CHEP/PCHS	
0461507	Oroville City Elementary	04615076003263	Ophir Elementary	
4469799	Pajaro Valley Unified	44697996049688	Hall (E. A.) Middle	
4469799	Pajaro Valley Unified	44697990105858	Pajaro Valley High	
4469799	Pajaro Valley Unified	44697996049811	Valencia Elementary	
1964857	Palmdale Elementary	19648576107635	Palm Tree Elementary	
1964857	Palmdale Elementary	19648576115273	Quail Valley Elementary	
1964865	Palos Verdes Peninsula Unified	19648656117584	Ridgecrest Intermediate	
1563362	Panama-Buena Vista Union	15633626115257	Reagan (Ronald) Elem.	
5475523	Porterville Unified	54755230116590	Harmony Magnet Academy	Charter
4570110	Redding Elementary	45701106050520	Sequoia Middle	
4570110	Redding Elementary	45701104530341	Stellar Secondary Charter High	Charter
1563578	Richland Union Elementary	15635780107771	Sequoia Elementary	
1075408	Riverdale Joint Unified	10754081035575	Riverdale High	
3367215	Riverside Unified	33672156032650	Jackson Elementary	
3367215	Riverside Unified	33672156032775	Victoria Elementary	
3175085	Rocklin Unified	31750856109870	Cobblestone Elementary	
3166928	Roseville Joint Union High	31669280116459	Antelope High	
3166928	Roseville Joint Union High	31669283136504	Roseville High	
2365607	Round Valley Unified	23656072334563	Round Valley High	
3467439	Sacramento City Unified	34674393434636	Johnson (Hiram W.) High	
3467439	Sacramento City Unified	34674396034169	Nicholas Elementary School	
2766167	San Antonio Union Elementary	27661676026629	San Antonio Elementary	
3667876	San Bernardino City Unified	36678763631090	Middle College	

Schools That Tested Students				
CD code	**District Name**	**CDS code**	**School Name**	**Charter/ Ind. Testing Charter?**
3667876	San Bernardino City Unified	36678763632809	Sierra High	
3610363	San Bernardino County Office of Education	36103630107466	Community/Independent Altern.	
3710371	San Diego County Office of Education	37103710120493	Monarch Elementary Community	
3710371	San Diego County Office of Education	37103710116038	North Region Court	
3768338	San Diego Unified	37683386039812	Keiller Leadership Academy	IT Charter
3868478	San Francisco Unified	38684786059869	Denman (James) Middle	
3868478	San Francisco Unified	38684786062061	Marina Middle	
3868478	San Francisco Unified	38684783830205	Wallenberg (Raoul) Traditional	
3467447	San Juan Unified	34674476034359	Arden Middle	
3467447	San Juan Unified	34674476097810	Woodside K–8	
4169047	San Mateo Union High	41690474130472	Burlingame High	
0761804	San Ramon Valley Unified	07618040736504	San Ramon Valley High	
1062414	Sanger Unified	10624141030766	Hallmark Charter	Charter
1062414	Sanger Unified	10624146007207	Washington Acad. Middle	
3066670	Santa Ana Unified	30666706058978	Lathrop (Julia C.) Intermediate	
3066670	Santa Ana Unified	30666700108365	Segerstrom High	
4269120	Santa Maria-Bonita	42691206045272	Bonita Elementary	
4169062	Sequoia Union High	41690624133716	Menlo-Atherton High	
3673957	Snowline Joint Unified	36739576112924	Heritage	
4169070	South San Francisco Unified	41690706059976	Alta Loma Middle	
1965037	South Whittier Elementary	19650376022834	Loma Vista Elementary	
5373833	Southern Trinity Joint Unified	53738335337423	Southern Trinity High	
5373833	Southern Trinity Joint Unified	53738336053805	Van Duzen Elementary	
1965045	Sulphur Springs Union	19650456022677	Canyon Springs Community Eleme	
5071290	Sylvan Union Elementary	50712900108761	Savage (Daniel J.) Middle	
5472249	Tulare Joint Union High	54722490116368	Mission Oak High	
0561580	Vallecito Union	05615806111884	Avery Middle	
5672652	Ventura Unified	56726525630793	Buena High	
3667918	Victor Elementary	36679186113757	Brentwood Elementary	
3667918	Victor Elementary	36679186037360	Park View Elementary	
3667934	Victor Valley Union High	36679343630761	Excelsior Education Center	IT Charter
3768452	Vista Unified	37684523730728	Rancho Buena Vista High	
3768452	Vista Unified	37684526040653	Vista Academy of Visual and P	
1973460	Walnut Valley Unified	19734606096358	Chaparral Middle	
1973460	Walnut Valley Unified	19734601939149	Walnut High	
5772694	Washington Unified	57726945735154	River City Senior High	
5772694	Washington Unified	57726946056402	Westmore Oaks Elementary	
0761796	West Contra Costa Unified	07617960732164	De Anza Senior High	
0761796	West Contra Costa Unified	07617966057210	Downer (Edward M.) Elementary	
0761796	West Contra Costa Unified	07617966005037	Washington Elementary	
1965094	West Covina Unified	19650941938513	Coronado High (Continuation)	
3066746	Westminster Elementary	30667466030761	Eastwood Elementary	
3066746	Westminster Elementary	30667466030928	Warner Middle	
1965128	Whittier Union High	19651281936889	Pioneer High	
1965136	William S. Hart Union High	19651360102475	West Ranch High	
5171464	Yuba City Unified	51714640107722	River Valley High	
3667959	Yucaipa-Calimesa Joint Unified	36679596037402	Calimesa Elementary	
3667959	Yucaipa-Calimesa Joint Unified	36679590117416	Chapman Heights Elementary	
3667959	Yucaipa-Calimesa Joint Unified	36679590101410	Oak View High School & Education	

Participating Districts That Used the CBT Tryout System and Uploaded Pre-ID Files

Participating Districts That Used CBT and Uploaded Pre-ID				
District code	**District Name**	**Students Tested**	**Students Pre-ID'd**	**Percent of Pre-ID'd Students that Tested**
1964212	ABC Unified	205	214	96%
0161119	Alameda City Unified	577	615	94%
1975713	Alhambra Unified	694	762	91%
3066423	Anaheim City	192	206	93%
3066431	Anaheim Union High	115	115	100%
1964246	Antelope Valley Union High	325	651	50%
3675077	Apple Valley Unified	117	138	85%
1964287	Baldwin Park Unified	579	609	95%
3667637	Bear Valley Unified	43	51	84%
1964303	Bellflower Unified	183	203	90%
4870524	Benicia Unified	68	137	50%
4970623	Bennett Valley Union Elementary	131	132	99%
1964311	Beverly Hills Unified	86	92	93%
1964329	Bonita Unified	302	305	99%
2765979	Bradley Union Elementary	8	8	100%
1363107	Calipatria Unified	45	45	100%
3066464	Capistrano Unified	357	369	97%
1964352	Centinela Valley Union High	32	32	100%
3066472	Centralia Elementary	57	58	98%
1964378	Charter Oak Unified	62	62	100%
2075606	Chawanakee Unified	134	216	62%
3667686	Colton Joint Unified	124	130	95%
1973437	Compton Unified	0	118	0%
1964444	Culver City Unified	457	470	97%
4275010	Cuyama Joint Unified	27	28	96%
5772678	Davis Joint Unified	33	38	87%
3768056	Del Mar Union Elementary	198	203	98%
0861820	Del Norte County Unified	84	214	39%
3367058	Desert Sands Unified	171	185	92%
5171373	East Nicolaus Joint Union High	15	39	38%
0961853	El Dorado Union High	4	4	100%
1964519	El Monte Union High	387	411	94%
3467314	Elk Grove Unified	338	377	90%
3768098	Escondido Union	302	305	99%
3667702	Etiwanda Elementary	190	194	98%
3166829	Eureka Union	161	168	96%
4369435	Evergreen Elementary	206	212	97%
4569989	Fall River Joint Unified	35	39	90%
3467330	Folsom-Cordova Unified	237	348	68%
3667710	Fontana Unified	388	729	53%
1010108	Fresno County Office of Education	10	15	67%
1062166	Fresno Unified	40	436	9%
1964550	Garvey Elementary	79	81	98%
4369484	Gilroy Unified	240	242	99%
1964568	Glendale Unified	316	317	100%
5071084	Gratton Elementary	6	6	100%

Participating Districts That Used CBT and Uploaded Pre-ID				
District code	**District Name**	**Students Tested**	**Students Pre-ID'd**	**Percent of Pre-ID'd Students that Tested**
4770326	Grenada Elementary	25	26	96%
3768130	Grossmont Union High	540	664	81%
1973445	Hacienda la Puente Unified	6	14	43%
1663925	Hanford Joint Union High	173	223	78%
3367082	Hemet Unified	235	273	86%
1964600	Hermosa Beach City Elementary	138	140	99%
5672462	Hueneme Elementary	120	122	98%
1363164	Imperial Unified	9	9	100%
3073650	Irvine Unified	236	256	92%
3367090	Jurupa Unified	317	444	71%
1073999	Kerman Unified	90	94	96%
1563529	Kern Union High	102	129	79%
1663941	Kings River-Hardwick Union Elementary	28	28	100%
5071142	Knights Ferry Elementary	12	13	92%
1563560	Lamont Elementary	131	134	98%
1864139	Lassen Union High	2	2	100%
2465730	Le Grand Union High	203	230	88%
2375218	Leggett Valley Unified	4	4	100%
1964725	Long Beach Unified	135	172	78%
3166845	Loomis Union Elementary	63	64	98%
1910199	Los Angeles County Office of Education	83	83	100%
1964733	Los Angeles Unified	164	5385	3%
1964774	Lynwood Unified	50	61	82%
3968593	Manteca Unified	171	532	32%
5872736	Marysville Joint Unified	241	274	88%
4373387	Milpitas Unified	167	183	91%
5071167	Modesto City Elementary	248	250	99%
1563677	Mojave Unified	71	135	53%
1062323	Monroe Elementary	21	22	95%
1964790	Monrovia Unified	3	3	100%
3367124	Moreno Valley Unified	370	395	94%
3667777	Morongo Unified	95	100	95%
1563685	Muroc Joint Unified	145	163	89%
3768221	National Elementary	62	65	95%
3475283	Natomas Unified	87	205	42%
2966357	Nevada Joint Union High	52	58	90%
3968627	New Jerusalem Elementary	200	318	63%
1964840	Norwalk-La Mirada Unified	34	34	100%
3667819	Ontario-Montclair Elementary	85	92	92%
3010306	Orange County Department of Education	17	17	100%
0461507	Oroville City Elementary	62	64	97%
4469799	Pajaro Valley Unified	440	558	79%
1964857	Palmdale Elementary	204	211	97%
1964865	Palos Verdes Peninsula Unified	191	197	97%
1563362	Panama-Buena Vista Union	144	149	97%
4970870	Piner-Olivet Union Elementary	0	23	0%
5475523	Porterville Unified	250	278	90%
4570110	Redding Elementary	231	249	93%
1563578	Richland Union Elementary	95	99	96%

Participating Districts That Used CBT and Uploaded Pre-ID				
District code	**District Name**	**Students Tested**	**Students Pre-ID'd**	**Percent of Pre-ID'd Students that Tested**
1075408	Riverdale Joint Unified	92	99	93%
3367215	Riverside Unified	183	188	97%
3175085	Rocklin Unified	55	63	87%
3166928	Roseville Joint Union High	145	147	99%
2365607	Round Valley Unified	32	34	94%
3467439	Sacramento City Unified	360	521	69%
2766167	San Antonio Union Elementary	19	19	100%
3667876	San Bernardino City Unified	192	279	69%
3610363	San Bernardino County Office of Education	12	16	75%
3710371	San Diego County Office of Education	15	31	48%
3768338	San Diego Unified	70	75	93%
3868478	San Francisco Unified	345	616	56%
3467447	San Juan Unified	208	211	99%
4169047	San Mateo Union High	89	93	96%
0761804	San Ramon Valley Unified	526	535	98%
1062414	Sanger Unified	419	479	87%
3066670	Santa Ana Unified	459	542	85%
4269120	Santa Maria-Bonita	21	60	35%
5472116	Sequoia Union Elementary	44	44	100%
4169062	Sequoia Union High	146	154	95%
3673957	Snowline Joint Unified	58	59	98%
4169070	South San Francisco Unified	234	257	91%
1965037	South Whittier Elementary	76	82	93%
5373833	Southern Trinity Joint Unified	16	19	84%
1965045	Sulphur Springs Union	73	73	100%
5071290	Sylvan Union Elementary	30	30	100%
5472249	Tulare Joint Union High	126	151	83%
0561580	Vallecito Union	23	26	88%
5672652	Ventura Unified	237	339	70%
3667918	Victor Elementary	198	231	86%
3667934	Victor Valley Union High	123	210	59%
3768452	Vista Unified	172	174	99%
1973460	Walnut Valley Unified	147	151	97%
5772694	Washington Unified	496	639	78%
0761796	West Contra Costa Unified	313	377	83%
1965094	West Covina Unified	68	86	79%
3066746	Westminster Elementary	248	259	96%
1965128	Whittier Union High	223	274	81%
1965136	William S. Hart Union High	62	66	94%
5171464	Yuba City Unified	47	52	90%
3667959	Yucaipa-Calimesa Joint Unified	159	212	75%

List of Schools/Districts Invited to Participate at the Time Recruiting Closed

Schools/Districts Invited to Participate			
District Name	**School Name**	**District Code**	**School Code**
ABC Unified	Bragg Elementary	1964212	6071369
ABC Unified	Cerritos High	1964212	1930056
Alameda City Unified	Alameda High	0161119	0130229
Alameda City Unified	Lincoln Middle	0161119	6090054
Alameda County Office of Education	FAME Public Charter	0110017	0109835
Alhambra Unified	Alhambra High	1975713	1930163
Alhambra Unified	Garfield Elementary	1975713	6011035
Alhambra Unified	Mark Keppel High	1975713	1934553
Anaheim City	Henry (Patrick) Elementary	3066423	6027387
Anaheim City	Orange Grove Elementary	3066423	0113712
Anaheim Union High	Western High	3066431	3038239
Anderson Union High	Anderson High	4569856	4530804
Anderson Valley Unified	Anderson Valley Jr./Sr. High	2365540	2330900
Antelope Valley Union High	Lancaster High	1964246	1995844
Apple Valley Unified	High Desert Premier Academy	3675077	3631009
Apple Valley Unified	Sandia Elementary	3675077	6107346
Atascadero Unified	Creston Elementary	4068700	6042915
Atascadero Unified	Santa Margarita Elementary	4068700	6042956
Baldwin Park Unified	Santa Fe Elementary	1964287	6110043
Baldwin Park Unified	Sierra Vista High	1964287	1938166
Bear Valley Unified	Big Bear Elementary	3667637	6105936
Bellflower Unified	Foster (Stephen) Elementary	1964303	6011704
Bellflower Unified	Ramona Elementary	1964303	6011696
Benicia Unified	Benicia High	4870524	4831004
Benicia Unified	Benicia Middle	4870524	6050983
Bennett Valley Union Elementary	Strawberry Elementary	4970623	6098248
Beverly Hills Unified	Horace Mann Elementary	1964311	6011779
Bonita Unified	San Dimas High	1964329	1937739
Bonsall Union Elementary	Bonsall West Elementary	3767975	0110122
Bradley Union Elementary	Bradley Elementary	2765979	6026017
Calipatria Unified	Young Middle	1363107	6008395
Capistrano Unified	Aliso Niguel High	3066464	3030574
Capistrano Unified	San Juan Hills High	3066464	0113381
Centinela Valley Union High	Leuzinger High	1964352	1935048
Centralia Elementary	Danbrook Elementary	3066472	6027684
Chaffey Joint Union High	Etiwanda High	3667652	3630373
Chaffey Joint Union High	Los Osos High	3667652	3631058
Chaffey Joint Union High	Montclair High	3667652	3633906
Chawanakee Unified	Minarets High	2075606	0117010
Chawanakee Unified	North Fork Elementary	2075606	6024079
Chawanakee Unified	O'Neals Digital Middle	2075606	0119826
Colton Joint Unified	Crestmore Elementary	3667686	6035604
Colton Joint Unified	Washington High	3667686	3630399
Compton Unified	Chavez (Cesar) High, Cont.	1973437	1995695
Corona-Norco Unified	Cesar Chavez Academy	3367033	6114151

Schools/Districts Invited to Participate			
District Name	**School Name**	**District Code**	**School Code**
Corona-Norco Unified	Norco High	3367033	3334232
Corona-Norco Unified	Rosa Parks Elementary	3367033	0113654
Culver City Unified	Culver City Middle	1964444	6057608
Cuyama Joint Unified	Cuyama Valley High	4275010	4231205
Davis Joint Unified	Da Vinci Charter Academy	5772678	0119578
Del Mar Union Elementary	Sage Canyon	3768056	6117923
Del Mar Union Elementary	Torrey Hills	3768056	6120596
Del Norte County Unified	Del Norte High	0861820	0833004
Del Norte County Unified	Sunset High	0861820	0836205
Delano Joint Union High	Delano High	1563412	1531672
Delhi Unified	Delhi High	2475366	2430114
Desert Sands Unified	Shadow Hills High	3367058	0118885
Durham Unified	Durham High	0461432	0433201
Earlimart Elementary	Earlimart CDS	5471902	0101303
East Nicolaus Joint Union High	East Nicolaus High	5171373	5132758
El Dorado Union High	Shenandoah High	0961853	0930214
El Monte Union High	El Monte High	1964519	1932664
Elk Grove Unified	Butler (Arthur C.) Elem.	3467314	6109516
Elk Grove Unified	Elk Grove High	3467314	3432572
Escondido Union	Mission Mid	3768098	6038210
Etiwanda Elementary	Day Creek Intermediate	3667702	0102947
Etiwanda Elementary	Heritage Intermediate	3667702	6119630
Eureka Union	Ridgeview Elementary	3166829	6111702
Evergreen Elementary	Cedar Grove Elementary	4369435	6067193
Evergreen Elementary	Whaley (O. B.) Elementary	4369435	6047161
Fall River Joint Unified	Fall River Junior-Senior High	4569989	4533600
Folsom-Cordova Unified	Cordova High	3467330	3431533
Folsom-Cordova Unified	Kitty Hawk	3467330	6033195
Fontana Unified	Jurupa Hills High	3667710	0120758
Fontana Unified	Locust Elementary	3667710	6106470
Fremont Union High	Homestead High	4369468	4333316
Fresno County Office of Education	Fresno County Court	1010108	1030337
Fresno Unified	Bullard Talent Project	1062166	6006100
Fresno Unified	Burroughs Elementary	1062166	6006118
Fresno Unified	Forkner Elementary	1062166	6101109
Fresno Unified	Powers-Ginsburg Elementary	1062166	6006431
Fresno Unified	School of Unlimited Learn	1062166	1030642
Fresno Unified	Winchell Elementary	1062166	6006571
Fruitvale Elementary	Endeavour Elementary	1563479	6112791
Galt Joint Union Elementary	Vernon E. Greer Elementary	3467348	0119420
Garvey Elementary	Bitely (Arlene) Elementary	1964550	6105886
Garvey Elementary	Monterey Vista Elementary	1964550	6013585
Garvey Elementary	Rice (Eldridge) Elementary	1964550	6013528
Garvey Elementary	Willard (Frances E.) Elementary	1964550	6013544
Gerber Union Elementary	Gerber Elementary	5271548	6053532
Gilroy Unified	South Valley Middle	4369484	6098214
Glendale Unified	Clark (A.W.) Magnet High	1964568	1996131
Glendale Unified	Verdugo Academy	1964568	1995497
Golden Plains Unified	Cantua Elementary	1075234	6005805

Schools/Districts Invited to Participate			
District Name	**School Name**	**District Code**	**School Code**
Golden Plains Unified	Rio Del Rey High (Continuation)	1075234	1030352
Gratton Elementary	Gratton Charter	5071084	0120089
Grenada Elementary	Grenada Elementary	4770326	6050777
Grossmont Union High	El Capitan High	3768130	3731809
Grossmont Union High	Grossmont Union High Special	3768130	3738077
Grossmont Union High	Monte Vista High	3768130	3734548
Hacienda la Puente Unified	Valley Alter. High, Cont.	1973445	1938935
Hanford Joint Union High	Hanford West High	1663925	1630169
Hemet Unified	Dartmouth Middle	3367082	6112007
Hermosa Beach City Elementary	Hermosa Valley Elementary	1964600	6095434
Hueneme Elementary	Hueneme Elementary	5672462	6055073
Hueneme Elementary	Richard Bard Elementary	5672462	6055107
Huntington Beach City Elementary	Dwyer (Ethel) Middle	3066530	6028849
Huntington Beach Union High	Edison High	3066548	3031895
Huntington Beach Union High	Marina High	3066548	3034410
Imperial Unified	Imperial Ave Holbrook High	1363164	1331115
Irvine Unified	Greentree Elementary	3073650	6089445
Irvine Unified	Meadow Park	3073650	6106850
Irvine Unified	Northwood Elementary	3073650	6100861
John Swett Unified	Carquinez Middle	0761697	6003685
Jurupa Unified	Rubidoux High	3367090	3337136
Jurupa Unified	Van Buren Elementary	3367090	6032239
Kerman Unified	Liberty Elementary	1073999	6119978
Kern Union High	Foothill High	1563529	1532605
Kern Union High	Highland High	1563529	1533330
Kings River-Hardwick Union Elementary	Kings River-Hardwick Elementary	1663941	6010474
Knights Ferry Elementary	Knights Ferry Elementary	5071142	6052609
La Canada Unified	La Canada Elementary	1964659	6014633
La Canada Unified	La Canada High	1964659	1934611
Lamont Elementary	Alicante Avenue Elementary	1563560	6009674
Lassen Union High	Diamond Mountain Charter High	1864139	0106385
Le Grand Union High	Le Grand High	2465730	2433001
Leggett Valley Unified	Leggett Valley High	2375218	2332724
Leggett Valley Unified	Whale Gulch High	2375218	2330207
Lindsay Unified	John J. Cairns Continuation	5471993	5430194
Live Oak Unified	Live Oak Alternative .	5171399	5130067
Long Beach Unified	Calif Acad Math and Science	1964725	1995539
Long Beach Unified	Renaissance High for the Arts	1964725	1996503
Loomis Union Elementary	Loomis Elementary	3166845	6031132
Los Angeles County Office of Education	Los Angeles International Cha	1910199	0109942
Los Angeles Unified	Addams (Jane) Cont. High	1964733	1932334
Los Angeles Unified	Aggeler Community Day	1964733	0102137
Los Angeles Unified	Alfred Bernhard Nobel Middle	1964733	6061543
Los Angeles Unified	Ann Street Elementary	1964733	6015812
Los Angeles Unified	Cal Burke High	1964733	0109512
Los Angeles Unified	Charles Maclay Middle	1964733	6058101
Los Angeles Unified	Elizabeth Learning Center	1964733	6016885
Los Angeles Unified	Foshay Learning Center	1964733	6061451
Los Angeles Unified	George S. Patton Continuation	1964733	1931518

Schools/Districts Invited to Participate			
District Name	**School Name**	**District Code**	**School Code**
Los Angeles Unified	George Washington Preparatory	1964733	1939305
Los Angeles Unified	Huntington Park Senior High	1964733	1934157
Los Angeles Unified	Ivy Academia	1964733	0106351
Los Angeles Unified	Los Angeles Senior High	1964733	1935352
Los Angeles Unified	Metropolitan Continuation	1964733	1931468
Los Angeles Unified	Olive Vista Middle	1964733	6061550
Los Angeles Unified	Orthopaedic Hospital	1964733	0106997
Los Angeles Unified	San Antonio Continuation	1964733	1930551
Los Angeles Unified	San Fernando Middle	1964733	6058283
Los Angeles Unified	Sierra Vista Elementary	1964733	6019228
Los Angeles Unified	Social Justice Leadership Acad	1964733	0122366
Los Angeles Unified	Ulysses S. Grant Senior High	1964733	1933795
Los Angeles Unified	View Park Continuation	1964733	1930429
Los Angeles Unified	Wilmington Middle	1964733	6058374
Lynwood Unified	Vista High (Continuation)	1964774	1935428
Manteca Unified	East Union High	3968593	3932001
Martinez Unified	Vicente Martinez High	0761739	0730440
Marysville Joint Unified	Lindhurst High	5872736	5830013
Marysville Joint Unified	Olivehurst Elementary	5872736	6056741
Merced County Office of Education	Valley Atwater Community Day	2410249	2430171
Merced Union High	Atwater High	2465789	2430601
Merced Union High	Independence High, Alt.	2465789	2430049
Merced Union High	Merced High	2465789	2435204
Milpitas Unified	Pomeroy (Marshall) Elementary	4373387	6047641
Milpitas Unified	Spangler (Anthony) Elementary	4373387	6047591
Modesto City Elementary	Franklin Elementary	5071167	6052690
Modesto City Elementary	Kirschen (Harriette) Elem	5071167	6105670
Modesto City Elementary	Sonoma Elementary	5071167	6093512
Mojave Unified	Hacienda Elementary	1563677	0113837
Monroe Elementary	Monroe Elementary	1062323	6006993
Monrovia Unified	Mountain Park	1964790	1996404
Moreno Valley Unified	Vista del Lago High	3367124	3331071
Morongo Unified	Twentynine Palms High	3667777	3636743
Muroc Joint Unified	Boron Junior-Senior High	1563685	1530997
Muroc Joint Unified	Desert Junior-Senior High	1563685	1531987
Muroc Joint Unified	West Boron Elementary	1563685	6009906
National Elementary	Rancho de la Nacion	3768221	6108559
Natomas Unified	Natomas Charter	3475283	3430659
Natomas Unified	Witter Ranch Elementary	3475283	0102277
Nevada Joint Union High	Sierra Mountain High	2966357	2930030
Nevada Joint Union High	William & Marian Ghidotti High	2966357	0112367
New Jerusalem Elementary	Delta Charter	3968627	6119309
Norwalk-La Mirada Unified	Chavez (Cesar) Elementary	1964840	6020853
Ontario-Montclair Elementary	El Camino Elementary	3667819	6036230
Ontario-Montclair Elementary	Howard Elementary	3667819	6036271
Orange County Department of Education	OCCS:CHEP/PCHS	3010306	3030632
Orchard Elementary	Orchard Elementary	4369633	6048185
Oroville City Elementary	Ophir Elementary	0461507	6003263
Pajaro Valley Unified	Hall (E. A.) Middle	4469799	6049688

Schools/Districts Invited to Participate			
District Name	**School Name**	**District Code**	**School Code**
Pajaro Valley Unified	Pajaro Valley High	4469799	0105858
Pajaro Valley Unified	Valencia Elementary	4469799	6049811
Palmdale Elementary	Palm Tree Elementary	1964857	6107635
Palmdale Elementary	Quail Valley Elementary	1964857	6115273
Palos Verdes Peninsula Unified	Ridgecrest Intermediate	1964865	6117584
Panama-Buena Vista Union	Reagan (Ronald) Elem.	1563362	6115257
Piner-Olivet Union Elementary	Northwest Prep at Piner-Olivet	4970870	0106344
Placer County Office of Education	iLearn Academy	3110314	0115675
Pomona Unified	Armstrong Elementary	1964907	6068787
Porterville Unified	Harmony Magnet Academy	5475523	0116590
Redding Elementary	Sequoia Middle	4570110	6050520
Redding Elementary	Stellar Secondary Charter High	4570110	4530341
Richland Union Elementary	Sequoia Elementary	1563578	0107771
Riverdale Joint Unified	Riverdale High	1075408	1035575
Riverside Unified	Jackson Elementary	3367215	6032650
Riverside Unified	Victoria Elementary	3367215	6032775
Rocklin Unified	Cobblestone Elementary	3175085	6109870
Roseville Joint Union High	Antelope High	3166928	0116459
Roseville Joint Union High	Roseville High	3166928	3136504
Round Valley Unified	Round Valley High	2365607	2334563
Rowland Unified	Yorbita Elementary	1973452	6022412
Sacramento City Unified	Bidwell (John) Elementary	3467439	6034045
Sacramento City Unified	Johnson (Hiram W.) High	3467439	3434636
Sacramento City Unified	New Technology High	3467439	0101881
Sacramento City Unified	Nicholas Elementary School	3467439	6034169
Sacramento City Unified	St. HOPE Public	3467439	0101048
San Antonio Union Elementary	San Antonio Elementary	2766167	6026629
San Bernardino City Unified	Middle College	3667876	3631090
San Bernardino City Unified	Sierra High	3667876	3632809
San Bernardino County Office of Education	Community/Independent Altern.	3610363	0107466
San Bruno Park Elementary	Belle Air Elementary	4169013	6044614
San Diego County Office of Education	Monarch Elementary Community	3710371	0120493
San Diego County Office of Education	North Region Court	3710371	0116038
San Diego Unified	Keiller Leadership Academy	3768338	6039812
San Francisco Unified	Denman (James) Middle	3868478	6059869
San Francisco Unified	Marina Middle	3868478	6062061
San Francisco Unified	Wallenberg (Raoul) Traditional	3868478	3830205
San Juan Unified	Arden Middle	3467447	6034359
San Juan Unified	Woodside K-8	3467447	6097810
San Mateo Union High	Burlingame High	4169047	4130472
San Ramon Valley Unified	San Ramon Valley High	0761804	0736504
San Ramon Valley Unified	Venture (Alternative)	0761804	0730218
Sanger Unified	Hallmark Charter	1062414	1030766
Sanger Unified	Washington Acad. Middle	1062414	6007207
Santa Ana Unified	El Sol Santa Ana Science	3066670	6119127
Santa Ana Unified	Lathrop (Julia C.) Intermediate	3066670	6058978
Santa Ana Unified	Segerstrom High	3066670	0108365
Santa Maria-Bonita	Bonita Elementary	4269120	6045272
Santa Rosa High	Carrillo (Maria) High	4970920	4930244

Schools/Districts Invited to Participate			
District Name	**School Name**	**District Code**	**School Code**
Selma Unified	Terry Elementary	1062430	6007322
Sequoia Union High	Menlo-Atherton High	4169062	4133716
Snowline Joint Unified	Heritage	3673957	6112924
South San Francisco Unified	Alta Loma Middle	4169070	6059976
South Whittier Elementary	Loma Vista Elementary	1965037	6022834
Southern Trinity Joint Unified	Southern Trinity High	5373833	5337423
Southern Trinity Joint Unified	Van Duzen Elementary	5373833	6053805
Stockton Unified	Edward C. Merlo Institute of	3968676	0115402
Stockton Unified	Stockton (Commodore) Skills	3968676	6098651
Stockton Unified	Stockton High	3968676	0119784
Stockton Unified	Taylor Leadership Academy	3968676	6042774
Stockton Unified	Wilson (Woodrow) Elementary	3968676	6042824
Sulphur Springs Union	Canyon Springs Community Eleme	1965045	6022677
Sylvan Union Elementary	Savage (Daniel J.) Middle	5071290	0108761
Tulare Joint Union High	Mission Oak High	5472249	0116368
Upper Lake Union High	Upper Lake Community Day	1764071	1730175
Vallecito Union	Avery Middle	0561580	6111884
Ventura Unified	Buena High	5672652	5630793
Ventura Unified	Buena Vista High	5672652	5630264
Victor Elementary	Brentwood Elementary	3667918	6113757
Victor Elementary	Park View Elementary	3667918	6037360
Victor Valley Union High	Excelsior Education Center	3667934	3630761
Vista Unified	Rancho Buena Vista High	3768452	3730728
Vista Unified	Vista Academy of Visual and P	3768452	6040653
Walnut Valley Unified	Chaparral Middle	1973460	6096358
Walnut Valley Unified	Walnut High	1973460	1939149
Washington Unified	River City Senior High	5772694	5735154
Washington Unified	West Sacramento Early College	5772694	0115329
Washington Unified	Westmore Oaks Elementary	5772694	6056402
West Contra Costa Unified	De Anza Senior High	0761796	0732164
West Contra Costa Unified	Downer (Edward M.) Elementary	0761796	6057210
West Contra Costa Unified	Lincoln Elementary	0761796	6004832
West Contra Costa Unified	Washington Elementary	0761796	6005037
West Covina Unified	Coronado High (Continuation)	1965094	1938513
Westminster Elementary	Eastwood Elementary	3066746	6030761
Westminster Elementary	Warner Middle	3066746	6030928
Westside Union Elementary	Quartz Hill Elementary	1965102	6023584
Whittier Union High	Frontier High, Cont.	1965128	1930338
Whittier Union High	Pioneer High	1965128	1936889
William S. Hart Union High	West Ranch High	1965136	0102475
Yuba City Unified	River Valley High	5171464	0107722
Yucaipa-Calimesa Joint Unified	Calimesa Elementary	3667959	6037402
Yucaipa-Calimesa Joint Unified	Chapman Heights Elementary	3667959	0117416
Yucaipa-Calimesa Joint Unified	Oak View High School & Education	3667959	0101410

Map of Invited Districts

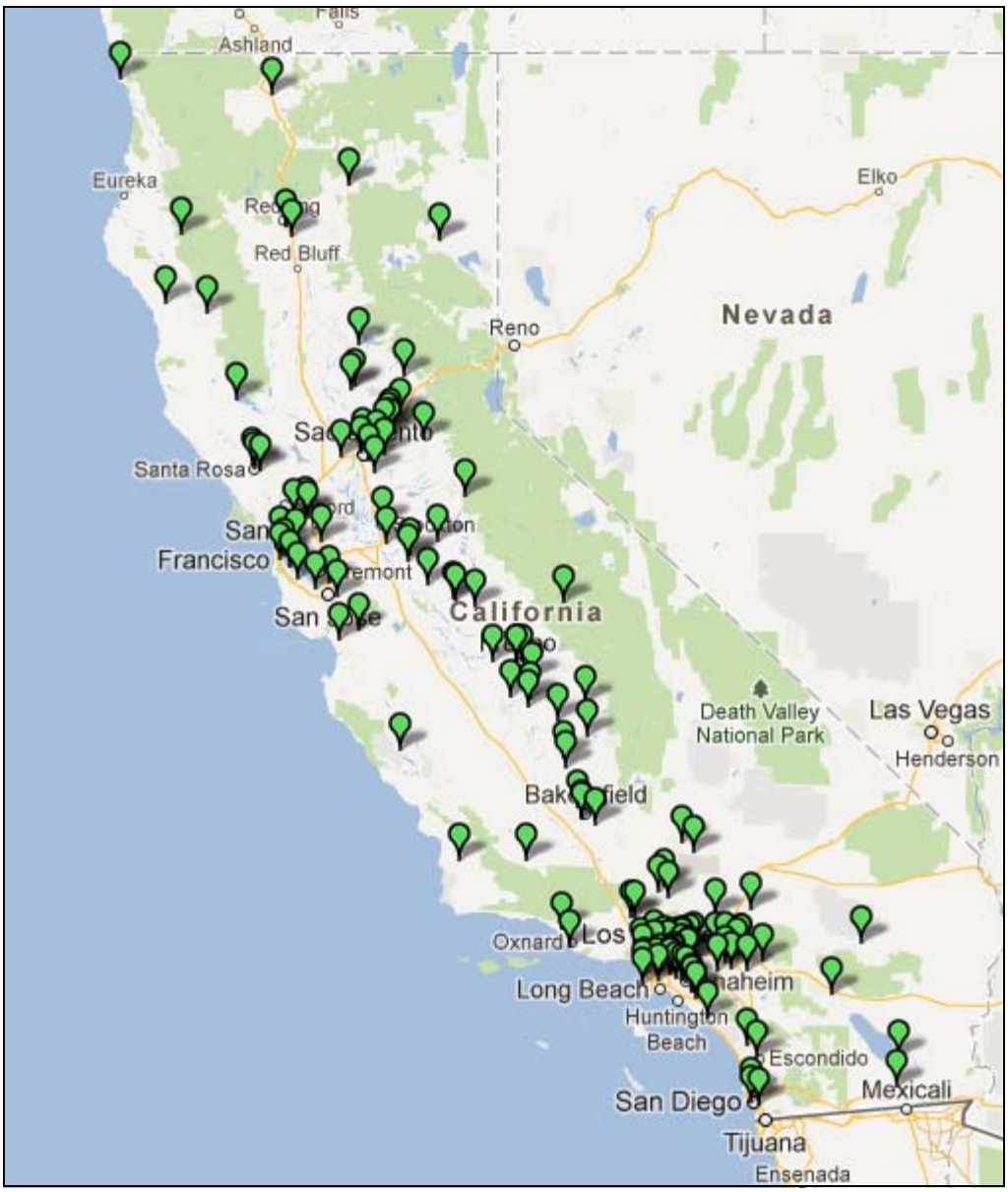


Figure A.1. Map of Invited Districts

Appendix B—Transcript of Survey Questions

Asterisks denote mandatory questions.

Questions for Students

Common. . .

Please mark only one option for each item.

1. How often do you use mobile electronic devices, such as e-reader (including Kindle, Nook, etc.), tablets (including iPad) or smart phones (including iPhone)?

❍ Five to seven days a week ❍ One to four days a week ❍ Less than one day a week ❍ Never

2. How often do you send/receive text messages?

❍ Five to seven days a week ❍ One to four days a week ❍ Less than one day a week ❍ Never

For all remaining questions please interpret the word computer to mean a desktop or a laptop.

3. How often did your science teacher use a computer in the class instruction last year?

❍ Five to seven days a week ❍ One to three days a week ❍ Less than one day a week ❍ Never

4. How much time did you spend doing work (including reading, writing and watching a video) for Science class on a computer last year? Include work you do in class and for homework.

❍ More than 10 hours a week ❍ 5 to 10 hours a week ❍ Less than 5 hours a week ❍ None

5. How often do you use a computer for all academic study at both school and home?

❍ Five to seven days a week ❍ One to four days a week ❍ Less than one day a week ❍ Never

6. How often do you use a computer to play games?

❍ Five to seven days a week ❍ One to four days a week ❍ Less than one day a week ❍ Never

7. How often do your other family members use computers at home?

❍ Five to seven days a week ❍ One to four days a week ❍ Less than one day a week ❍ Never

8. How would you rate your computer skills?

❍ Expert ❍ Pretty good ❍ So-so ❍ Just learning

9. I am relaxed when I am working on a computer.

❍ Strongly agree ❍ Agree ❍ Disagree ❍ Strongly disagree

10. What types of computer do you feel comfortable using?

❍ PC ❍ Mac ❍ Both PC and Mac ❍ Neither PC nor Mac

11. Have you ever taken a school test on a computer before? The test may have been a classroom test, a standardized test, a practice test or quiz, or any other type of test.

❍ Yes, often ❍ Yes, a few times ❍ Yes, once or twice before ❍ No, never before

12. I can type as quickly and accurately on a computer test as I write on a paper and pencil test.

❍ Strongly agree ❍ Agree ❍ Disagree ❍ Strongly disagree

13. Would you rather take tests on a computer or on paper?

❍ Computer ❍ Paper

If you answered "Computer," answer #14 below and skip #15. If you answered "Paper," skip #14 and answer #15.

14. Check the reason(s) if you prefer taking a test on computer (Check all the options that apply to you)

❑ I have a lot of practice in taking tests on computer

❑ Testing on the computer is like completing my class assignments

❑ I can write better and faster on computer than on paper

❑ I can focus better on a computer

❑ Test questions are easier to read on computer than on paper

❑ I can better show what I know in science.

❑ Taking a test on computer is more fun than taking it on paper

15. Check the reason(s) if you prefer taking a test on paper (Check all the options that apply to you).

❑ I haven't had enough practice in taking tests on computer

❑ Testing with paper and pencil is like completing my class assignments

❑ I can write better and faster on paper than on computer

❑ It is harder to focus while taking a test on computer than on paper

❑ Test questions are easier to read on paper than on computer

❑ School computers are not like the computers I used at home

❑ I worry about making a mistake when I take a test on a computer

. . . Grade Six Only

Questions #16–19 specifically relate to your experience investigating stream speed in the watershed simulation.

16. The instructions for completing this task were clear and easy to understand.

❍ Strongly agree ❍ Agree ❍ Disagree ❍ Strongly disagree

17. I could understand what the pictures were showing.

❍ Strongly agree ❍ Agree ❍ Disagree ❍ Strongly disagree

18. It was easy to run the animation to collect data.

❍ Strongly agree ❍ Agree ❍ Disagree ❍ Strongly disagree

19. Plotting the data points on the graph was the easiest part of the Steam Speed investigation.

❍ Strongly agree ❍ Agree ❍ Disagree ❍ Strongly disagree

Questions #20–23 specifically relate to your experience investigating water moving through soil in the watershed simulation.

20. The instructions for completing this task were clear and easy to understand.

❍ Strongly agree ❍ Agree ❍ Disagree ❍ Strongly disagree

21. I could understand what the pictures were showing.

❍ Strongly agree ❍ Agree ❍ Disagree ❍ Strongly disagree

22. Running the animation to collect data was easy to do.

❍ Strongly agree ❍ Agree ❍ Disagree ❍ Strongly disagree

23. I understood what the data in the table meant compared to the animation.

❍ Strongly agree ❍ Agree ❍ Disagree ❍ Strongly disagree

. . . Grade Nine Only

Questions #16–19 specifically relate to your experience with the introduction to the functions in the simulation of the toy car.

16. The instructions for completing this task were clear and easy to understand.

❍ Strongly agree ❍ Agree ❍ Disagree ❍ Strongly disagree

17. I could understand what the pictures were showing.

❍ Strongly agree ❍ Agree ❍ Disagree ❍ Strongly disagree

18. Dragging the car was easier than using the timer to collect data.

❍ Strongly agree ❍ Agree ❍ Disagree ❍ Strongly disagree

19. I understood how the data from the timer related to the increments on the track.

❍ Strongly agree ❍ Agree ❍ Disagree ❍ Strongly disagree

Questions #20–23 specifically relate to your experience with the toy car simulation.

20. The instructions for each item in the simulation were clear and easy to understand.

❍ Strongly agree ❍ Agree ❍ Disagree ❍ Strongly disagree

21. I could understand what the highlighting on the track meant during the simulation.

❍ Strongly agree ❍ Agree ❍ Disagree ❍ Strongly disagree

22. I was able to easily observe the timer and the motion of the car at the same time.

❍ Strongly agree ❍ Agree ❍ Disagree ❍ Strongly disagree

23. It was easy to work with more than one item on the screen at one time.

❍ Strongly agree ❍ Agree ❍ Disagree ❍ Strongly disagree

. . . Biology Only

Questions #16-19 specifically relate to your experience generating a Punnett Square in the simulation of fruit fly genetics.

16. The instructions for completing this task were clear and easy to understand.

❍ Strongly agree ❍ Agree ❍ Disagree ❍ Strongly disagree

17. I could understand what the pictures were showing.

❍ Strongly agree ❍ Agree ❍ Disagree ❍ Strongly disagree

18. It was easy to understand how to use the drop down lists.

❍ Strongly agree ❍ Agree ❍ Disagree ❍ Strongly disagree

19. I understood how to drag and drop images to complete the task.

❍ Strongly agree ❍ Agree ❍ Disagree ❍ Strongly disagree

Questions #20–23 specifically relate to your experience creating a pie graph in the simulation of fruit fly genetics.

20. The instructions for completing this task were clear and easy to understand.

❍ Strongly agree ❍ Agree ❍ Disagree ❍ Strongly disagree

21. I could understand what the pictures were showing.

❍ Strongly agree ❍ Agree ❍ Disagree ❍ Strongly disagree

22. It was easy to draw lines on the pie chart to create my graph of the data.

❍ Strongly agree ❍ Agree ❍ Disagree ❍ Strongly disagree

23. I understood what the pie graph was meant to represent in the task.

❍ Strongly agree ❍ Agree ❍ Disagree ❍ Strongly disagree

Questions for Test Administrators

1. How did your students react to the STAR CBT testing experience?

❍ Positive ❍ Indifferent

❍ Negative ❍ Other:

Additional comments regarding students:

2. Were the Directions for Administration helpful?

❍ Yes, very helpful ❍ Somewhat helpful ❍ Not helpful ❍ Didn't use

Additional comments or suggestions for improvements to the DFAs:*

3. Did your students use scratch paper?*

❍ No ❍ Yes, a few ❍ Yes, many ❍ Yes, all ❍ I didn't distribute scratch paper

4. Was the "Print Tickets" function easy to use?

❍ Yes, easy ❍ Somewhat easy ❍ Not easy ❍ Didn't use

5. Did you print one ticket per page, or twelve?

❍ One ticket per page ❍ Two tickets per page

Additional comments or suggestions for improvements to the ticketing function:*

6. During the past school year, how many hours of instructional technology professional development did teachers at your school complete, on average?

❍ More than 20 hours ❍ Between 12 and 20 hours ❍ Between 4 and 12 hours ❍ Less than 4 hours

7. Estimate what percentage of teachers in your school had experience with computer-based assessment for the subjects they teach.

❍ More than 80% ❍ Between 50% and 80% ❍ Between 20% and 50% ❍ Below 20%

8. How many technology experts provide informal professional development in your school (including support such as coaching, mentoring, and co-teaching)?

Number of technology experts

How many Science teachers at your school?

Number of Science teachers

9. What percentage of Science teachers at your school use computers at least once a week as part of their classroom instruction?

❍ More than 80% ❍ Between 50% and 80% ❍ Between 20% and 50% ❍ Below 20%

10. Estimate the percentage of Science teachers at your school who assign homework at least once a week that requires the use of computers (e.g., research, simulation, multimedia presentation, and online collaboration)?

❍ More than 80% ❍ Between 50% and 80% ❍ Between 20% and 50% ❍ Below 20%

11. Estimate the percentage of Science teachers at your school use computers no more than once a month as part of classroom instruction?

❍ More than 80% ❍ Between 50% and 80% ❍ Between 20% and 50% ❍ Below 20%

12. During the past school year, how often did teachers at your school use instructional technology with students for activities such as research, multimedia, simulations, data interpretation, communications, and collaboration?

❍ Nearly every day ❍ About once a week ❍ About once a month ❍ Rarely or never

13. How prepared is your school for administering computer-based tests? (including infrastructure, computer assessment software, administrators', teachers' and students' technology readiness)

❍ Very well prepared ❍ Somewhat prepared ❍ Unprepared ❍ Very much unprepared

14. Did your school have experience administering computer-based tests before this tryout?

❍ Yes ❍ No

Check all that apply:

❑ District Benchmarks ❑ Teacher Classroom Tests

15. What is the total number of computers (including desktops, laptops and tablets) in your school that can be used for educational purposes by students?

Number of students in your school:

Any additional comments or suggestions regarding the STAR CBT Tryout?

General Questions for All LEA and School Staff

What was your role in administering the STAR CBT Tryout?

- ❑ STAR Coordinator for your LEA
- ❑ Technology Coordinator for your LEA
- ❑ Test Administrator (any involvement with testing at the school level)
- ❑ Other: [ ]

(Note: select more than one if you filled multiple roles)

Did you find the STAR CBT Tryout system easy or difficult to use?

❍ Very Easy ❍ Easy ❍ Neutral ❍ Difficult ❍ Very Difficult

Did the STAR CBT Tryout technology work as you expected?

❍ Yes ❍ No

If not, why?

- ❑ Problems with school equipment, technology, or software
- ❑ Lack of resources to conduct testing
- ❑ System was difficult to use
- ❑ Training materials were inadequate
- ❑ Other: [ ]

If you experienced any problems with STAR CBT, please list them here:

[ ]

Which of the following resources did you use and how helpful were they?

Webcast

❍ Very helpful ❍ Helpful ❍ Somewhat helpful
❍ Not very helpful ❍ Not at all helpful ❍ Did not use

Manual

❍ Very helpful ❍ Helpful ❍ Somewhat helpful
❍ Not very helpful ❍ Not at all helpful ❍ Did not use

Tutorial

❍ Very helpful ❍ Helpful ❍ Somewhat helpful
❍ Not very helpful ❍ Not at all helpful ❍ Did not use

Management Tools

❍ Very helpful ❍ Helpful ❍ Somewhat helpful
❍ Not very helpful ❍ Not at all helpful ❍ Did not use

STAR Technical Assistance Center

❍ Very helpful ❍ Helpful ❍ Somewhat helpful

❍ Not very helpful ❍ Not at all helpful ❍ Did not use

Additional comments regarding any of the above:

Questions for STAR Coordinators

Were your school(s) able to successfully administer the CBT tests?

❍ Yes ❍ Some schools yes, some no ❍ No

Please select the reason(s) your schools were unable to test students.

❍ Didn't have the right technology in place
❍ No staff time available to support testing
❍ School schedules conflicted with October 1–12 testing window
❍ Other:

Is there any other information that would help us understand why your schools could not test?

After participating in the STAR CBT Tryout, have your perceptions of your LEA's preparedness for CBT changed?

❍ We are better prepared than I thought ❍ We are less prepared than I thought
❍ We are about as prepared as I thought

Any additional thoughts about your LEA's readiness for CBT?

What advice do you have for schools and districts that did not participate, to help them prepare for CBT?

If given the choice to administer future STAR tests on paper or using the CBT Tryout system, what would you choose?

❍ Computer ❍ Paper ❍ A mix of computer and paper

Roughly how many students might you prefer to test on computer in your first CBT administration?

❍ 10% ❍ 25% ❍ 50% ❍ 75%

How easy was it to manage the following aspects of CBT Testing?

Use and distribute passwords:

❍ Very Easy ❍ Easy ❍ Neutral ❍ Difficult ❍ Very Difficult ❍ Not applicable

Schedule testing sessions:

❍ Very Easy ❍ Easy ❍ Neutral ❍ Difficult ❍ Very Difficult ❍ Not applicable

Generate and upload Pre-ID files:

❍ Very Easy ❍ Easy ❍ Neutral ❍ Difficult ❍ Very Difficult ❍ Not applicable

Monitor students and see results:

❍ Very Easy ❍ Easy ❍ Neutral ❍ Difficult ❍ Very Difficult ❍ Not applicable

Who was responsible for Pre-ID in your district?

❍ LEA STAR Coordinator ❍ School Administrators ❍ Both

Additional comments regarding any of the above:

Questions for Technology Coordinators

How did you install the testing system?

- ❑ Went to each school and installed on every computer
- ❑ Installed remotely
- ❑ Other:

If you had any problems with the installation of the testing system or the LCS system, please describe them here:

What technical support did you provide to schools during testing?

Appendix C—Additional Analyses Based on Survey Responses and Site Visits

Site Visits

Table C.1. Sites Visited

School	School District	Grade
Strawberry Elementary	Bennet Valley Union	6
California Academy of Mathematics and Science	Long Beach Unified	9
Cobblestone Elementary	Rocklin Unified	6
Cordova High	Folsom Cordova Unified	Biology
DeAnza High	West Contra Costa Unified	Biology
Elk Grove High	Elk Grove Unified	Bio, part 2
Howard Elementary	Ontario-Montclair	6
Loma Vista Elementary	South Whittier	6
Meadow Park Elementary	Irvine Unified	6
Ridgecrest Intermediate	Palos Verdes Peninsula Unified	6
Ridgeview Elementary School	Eureka Union	6
San Ramon Valley High	San Ramon Valley Unified	9
Witter Ranch Elementary	Natomas Unified	6

ETS staff visited 13 different sites during testing; sites are shown in Table C.1, above. In one case, two observers visited the same site. Observers viewed between one and five testing sessions at each site. Seven of the sites were in southern California and six were in northern California. Among the thirteen observed sites, eight tested in grade six, two tested students in grade nine, and three tested for Biology.

This section reviews answers to questions and additional observations related to problems and suggestions. Note that comments in a bulleted list are taken directly from the survey responses.

Paper Testing vs. CBT Testing

Eight administrators were asked whether they thought CBT testing was easier to administer than paper-based testing. Virtually all adults managing the testing at schools said CBT testing was easier, provided they received adequate technical support. According to feedback from these administrators, the primary advantage of CBT over paper-based testing is the greatly reduced paper-handling workload.

In some cases, the administrators running the tests had no previous experience with STAR testing, and thus had no basis for comparison. These administrators included technology experts and a physical education teacher. More typical was the teacher who was enthusiastic that after testing she could just collect the tickets and scratch paper and not have additional tasks related to organizing student answer documents for someone else to pack and ship. Another teacher simply stated that it was very easy to administer the CBT test.

Training and Materials

At nine sites, administrators provided information on training and materials. The *CBT Tryout Manual* was most commonly used, and administrators provided many suggestions for improving it (see page 97 for the responses). Another popular resource was the Webcast. In some cases, administrators received no training. Table C.2 on the next page shows the number who said they used a particular

training material. The numbers add to more than nine since at some locations, more than one person responded.

Table C.2. Use of Training Materials

Training used	Number who said they used it
Manual	6
Webcast	6
FAQs	5
Checklists	4
Training videos	4
No training	3

Staff at testing sites explained that making teachers comfortable with CBT goes beyond training materials. They need to acquire familiarity with processes with hands-on experience. One observer quoted an administrator:

> *"Most of the test administrators/teachers would need to work to a certain comfort level doing a CBT and getting used to whatever new processes are in place both for the assessment program and district IT requirements."*

Directions for Administration (DFA)

Eleven of the observers indicated whether or not the teacher administering the test used the *DFA* (CDE, 2012a). Nine teachers used the *DFA* verbatim. One did not use it at all, and another just referenced the *DFA* in general. Most test administrators agreed that the *DFA* was adequate. A few teachers mentioned they liked the simplicity of the directions, although another said there was too little information. The biggest complaint was about the lack of information about the transition between Part 1 and Part 2.

Test Session Tickets

Of the nine observers who mentioned tickets, only two mentioned having any problems. A few administrators said they wished they had used the optional function for grouping students within a school by testing time and location, and some mentioned this function needed to be more user-friendly. When asked whether they opted to print tickets one-per-page or twelve-per-page, all but two printed one ticket per page, preferring the convenience of not needing to cut out the tickets over paper-saving. One-page tickets also doubled as scratch paper available to students, which was a testing requirement. Test administrators reported that not many students used the scratch paper.

Student Reactions

Students overwhelmingly preferred CBT over paper testing. They were more engaged, found it easier, less distracting, less confusing, and more interesting. Students typically took less time to complete the tests than anticipated. What follows is a sampling of student statements about CBT that were recorded during site visits:

> *"It was easier than taking a paper test. On paper it is easy to get distracted with all the graphs and charts, but on this test I was able to focus on each question."*
>
> *"Answering the questions was less confusing than the regular STAR test."*
>
> *"The difficulty of the content was the same, but taking the test was easier."*
>
> *"Building the graphs and the punnett square was easier and better than answering (MC) questions about them."*

"I liked being able to type in an answer."

"The questions were easier to see. I liked being able to make the pictures bigger so I could see them more clearly."

"The test was more interesting with color and different ways of answering questions."

In one case, the instructor projected the tutorial to save time. Students liked the practice test, but there were complaints about the tutorial. Students often failed to read the instructions at the top or to know to click on the orange dot. See the "Tutorials" topic in the "Suggestions for Improvement" section for suggestions for improvement.

One administrator comment about the students' interaction with the CBT (at a continuation school) was that "Once the novelty wore off and the content became challenging, students disengaged and rushed to finish." Another student said she just didn't like computers. In one case, students were given Part 1 and Part 2 back to back, and some students seemed tired at the end, going slower and yawning.

Only a few students used the striker, highlighter, and other tools. One student said, "The tools helped me get through the test faster."

Suggestions for Improvement

During site visits, test administrators offered several suggestions to improve CBT Testing related to the transition between Part 1 and Part 2 of the tests, technical issues, the manual, the *DFA*, the tutorial, transition to "live" CBT, and general suggestions. The most frequent suggestions were related to the Part 1/Part 2 transition issue.

Part 1 to Part 2: Continue or Exit

It was recommended that the local technical person have the ability to turn the "continue test" function off or on so students could not continue to Part 2 if it was intended to be administered at a later time. Another suggestion was to include a big red "stop" sign, similar to paper-based STAR tests. This topic should be covered in grade-level appropriate detail in the manual and the *DFA*.

Technical

- Ensure that the technology coordinator is available during testing, ideally, on site.
- Explain what to do about the modules/icons after testing.
- On Pre-ID, make birth date a single field rather than three.

Manual

- Explain in greater detail the student "Reactivate" function.
- Include a troubleshooting section that includes, for example, what to do if a student's computer fails during testing.
- For tickets, give a more detailed explanation of the group sorting feature.

DFA

- Include screen shots.

Tutorial

- Make the instructions text a larger font.
- Gray out computer screen examples, so it is clear they are nonfunctional for tutorial purposes.

General suggestions

- Revise how testing time is estimated. Indicate the average time the test should take, along with minimum and maximum so schools can better schedule. [**Note:** This was not possible for the CBT Tryout, since this was the first time students were taking these test forms and no previous knowledge of testing times was available.]
- Don't separate Biology by grade. This created too much extra work.

Analyses of Post-Test Survey Open-ended Responses

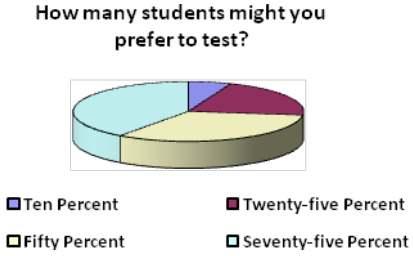
Aggregated Responses

Responders were asked to categorize their role in administering the CBT Tryout based on the three roles that were defined in the training materials. Of the 243 responses, 29 percent were STAR coordinators for their LEA, 14 percent were technology coordinators for their LEA, 33 percent were test administrators at the school level, and 15 percent filled multiple roles or did not state their role. The following are summary results based on all responses Comments in a bulleted list are taken directly from the survey responses.

Summarized responses are based on all 243 surveys submitted. The respondents represent 86,107 students who have 12,500 computers available for educational purposes at their schools with a ratio of 1 computer per 15 students.

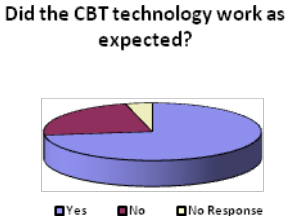
Did you find the STAR CBT Tryout system easy or difficult to use?

16%	Very Easy
55%	Easy
21%	Neutral
6%	Difficult
1%	Very Difficult
3%	No response



Did the STAR CBT Tryout technology work as you expected?

73%	Yes
23%	No
4%	No response



Of the 22 percent reporting the CBT Tryout technology did not work as expected, two-thirds cited problems with school equipment, technology, or software as the main reason.

Which of the following resources did you use and how helpful were they?

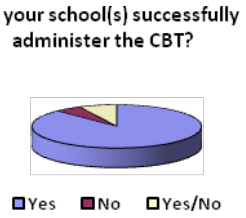
	Very helpful	Helpful	Somewhat helpful	Not very helpful	Not at all helpful	Did not use/ No response
Webcast	15%	30%	14%	6%	3%	33%
Manual	17%	35%	21%	1%	0%	25%
Tutorial	30%	36%	14%	1%	0%	19%

	Very helpful	Helpful	Somewhat helpful	Not very helpful	Not at all helpful	Did not use/ No response
Management Tools	24%	41%	10%	3%	0%	21%
STAR TAC	21%	16%	6%	2%	0%	53%

Were your schools able to successfully administer the CBT tests?

This question was directed only to LEA STAR coordinators.

85% Yes
5% No
7% Some yes, some no



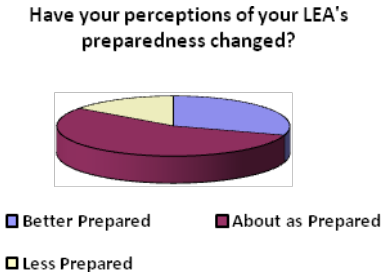
Of the 5 percent that said no, the following reasons were cited:

- For us, I believe it was the late notification of my district's ability to participate which was the most challenging.
- Didn't have the right technology in place
- Program did not work
- IP address failure to connect to server
- No staff time available to support testing
- Principal did not want to participate in another field test.

After participating in the STAR CBT Tryout, have your perceptions of your LEA's preparedness for CBT changed?

This question was directed to LEA STAR coordinators.

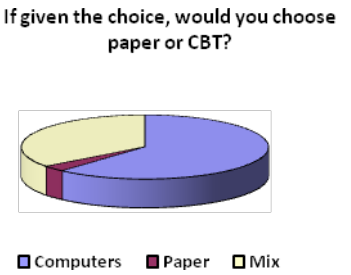
30% We are better prepared than I thought
55% We are about as prepared as I thought
15% We are less prepared than I thought



If given the choice to administer future STAR tests on paper or using the CBT Tryout system, which would you choose?

This question was directed to LEA STAR coordinators.

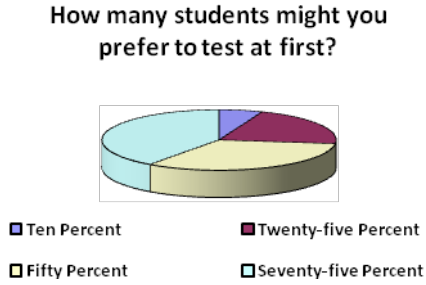
61%	Computer
3%	Paper
35%	A mix of computer and paper



Roughly how many students might you prefer to test on computer in your first CBT administration?

This question was answered by LEA STAR coordinators who responded "computer" or "mix of computer and paper" to the previous question.

6%	Ten percent of students
21%	Twenty-five percent of students
33%	Fifty percent of students
40%	Seventy-five percent of students



Responses from STAR Coordinators

Seventy-three district STAR coordinators responded to the survey after testing. What follows is an analysis and summary of the open-ended responses provided by these STAR coordinators.

Why didn't the STAR CBT Tryout work as you expected?

Most coordinators reported that the CBT Tryout technology did work as expected. However, 25 of the responding coordinators said the CBT did not work as expected. All but two of the reasons were due to technical problems. The reasons for this response included the following:

- The system was difficult (four mentioned this)
- Lack of resources (three mentioned this)
- Difficulty with pausing (three mentioned this)
- Software
- Installation
- Having to change IP addresses on each computer daily
- Screen freezes
- Did not work well on Mac

Problems experienced

Fifty-two STAR coordinators indicated having some degree of problems with the CBT Tryout. Of these, forty-three reported technical problems. The most frequently mentioned problems had to do with

transitioning from Part 1 to Part 2 and a related problem with pausing and reactivation of student testing sessions. Coordinators said the directions for Part 2 were not explicit enough. Coordinators cited difficulty in reactivating those students' files and also reactivating students who had selected "Pause."

Additional issues that were mentioned are as follows:

- Server problems, both local and offsite, that inhibited dragging and dropping and booted students from the system (six mentioned this)
- The slow speed of the system (five mentioned this)
- Difficulty with installation; some had to update Adobe or install Windows XP (four mentioned this)
- Difficulty logging on (four mentioned this)
- Lack of resources, including computers, technical help, Internet and Wi-Fi connections (three mentioned this)

Webcast, manual, tutorial, management tools, and the STAR Technical Assistance Center

Coordinators were asked about the Webcast, manual, tutorial, management tools, and STAR TAC. Twenty-five coordinators added open-ended comments to their responses:

- Webcast: The two respondents complained that it froze and was too long.
- Tutorial: One said it was useful while another stated that it was confusing.
- Technical support: Twelve mentioned this in their response, with ten giving praise for the support. Two responses mentioned that they couldn't get through by phone.

Readiness for testing

Forty-three coordinators provided open-ended responses regarding their LEAs' readiness for CBT. Nine said they were ready, while 34 indicated that additional resources were needed in order to test the entire district on all tests via CBT.

Positive comments were as follows:

- We were better prepared than the teachers thought.
- We're ready for all grades.
- CBT helped us see what we need to do to have smooth testing.
- This was a good opportunity to view CBT.
- CBT worked well on our new computers; I'd like to see how the old computers could handle it.
- We want to be part of future pilots.

Twenty, or 47 percent, said they do not have enough computers. Other comments about computers related to issues with wireless systems and the need for newer computers. Further comments were as follows:

- The installation was a burden.
- Need to be able operate off a server for the whole school.
- Computer network concerns.

Seven, or 16 percent, reported a lack of technical assistance in their school district. Others cited scheduling problems, particularly with minimum days and block schedules.

Additional information about passwords, scheduling, Pre-ID, monitoring students

Forty coordinators, or 93 percent, made additional comments:

- Scheduling: Three of the four who discussed scheduling said it was easy, but would be difficult if all students were to test. Another said that with 50-minute classes, students had to go to the computer lab at least three times to take the tutorial, practice test, and test.
- Pre-ID: Twelve mentioned Pre-ID and nine of them said they reached out for technical help. Additional problems and suggestions were as follows:
  - Student names were limited to eleven characters; more characters needed
  - We should be able to download the data from CALPADS.
  - Optional test session grouping function should accept numbers in the "location" field.
- Positive remarks: Eight coordinators gave positive comments, including liking no paper, that students were engaged, and that it was easy to use. One coordinator quoted a student as saying: "At least I don't fall asleep with this test."
- Other: Fourteen district STAR coordinators wrote comments in this section that were more general than the listed categories.
  - Three expressed concerns about the number of computers that would be needed for 100 percent CBT for the entire district.
  - Two pointed out the need to randomize the questions or have different versions to prevent cheating.
  - One said the school needs more time to get ready.
  - One noted that CBT will be difficult for younger students.
  - One coordinator gave this list of what students did not like: (1) Prefer headphones to keep them focused; (2) Looking down at paper helps them concentrate more than looking ahead at screen; (3) Looking at screen for that long hurts their eyes; (4) Font size too small.

Additional comments regarding students

Seven coordinators, or 16 percent, included more information about students. Six made positive remarks about how much students liked engaging in the CBT and saving trees. Three of them gave some negative responses about the computer crashing and students not knowing how to progress through the tutorial and practice tests.

Additional suggestions for DFA

Ten coordinators, or 23 percent, offered suggestions for the *DFA*. Two of them liked that the *DFA* was simple and helpful. Suggestions included:

- Give more specific technical information for teachers.
- Add more explanation of the Part 1/Part 2 transition. (requested by two coordinators)
- Add rationale for each style of question.

Additional comments or suggestions regarding the STAR CBT Tryout

Twenty-eight coordinators, or 65 percent, made general comments about the CBT Tryout. Ten coordinators expressed a concern that they do not have enough computers. Six coordinators made positive remarks about the following:

- Needing less time for CBT than paper-based tests
- The ease of monitoring

- The value of instant results
- The fact that students like CBT so much

Other comments and suggestions included the following:

- Move the "Exit" button to the upper right
- Students are not prepared to use the resources or the tools
- Need more time for technical preparation
- Randomize the questions to prevent cheating
- There's no space at workstations in the lab for scratch paper

Responses from Test Administrators

Eighty-five respondents were involved in administering the test at schools. Eighty-one were school administrators and four were district coordinators who were also involved with testing at the school level.

Most administrators reported some degree of technical difficulty with the CBT Tryout. A lack of resources or technical support was a common technical issue. Some reported that students progressed from Part 1 to Part 2 of the test without stopping, as was intended by the administrator. Other administrators mentioned problems with computers freezing, problems with drag-and-drop functionality in some items, difficulty in logging on, installing the system, and with school Wi-Fi systems.

Additional comments on Webcast, manual, tutorial, Management Tools, STAR TAC

- Webcast stopping or not being useful (three mentioned)
- STAR TAC was very helpful (three mentioned)
- Tutorial did not address all item types
- Need to provide student scores based on standards and item type
- DFA needs more explanation of how to reactivate student test sessions and on the transition from Part 1 to Part 2 of the test.
- Would like to know student results.
- Appreciated having a STAR representative on site during testing.

Tips for schools and districts that did not participate

- Be sure administrators are familiar with Wi-Fi and Wi-Fi configurations.
- Participate in a pilot or tryout.
- Provide technical assistance at the school on the day of testing.

Additional comments on passwords, scheduling, Pre-ID, monitoring, and seeing results

- It's easy to schedule and monitor students.
- You should be able to upload Pre-ID info directly from CALPADS.
- Technical coordinator handled Pre-ID.

How did you install the testing system?

Four test administrators, or 5 percent, mentioned installing the testing system. All reported installing the system on each computer individually rather than taking advantage of auto-install features.

- Installation was labor-intensive.
- Installation took several tries due to non-uniformity of the age of our systems, the design of the labs, and the damage done in a lightning strike a few weeks ago.

Additional comments about students' reactions to CBT

Thirty-three test administrators, or 39 percent, gave additional comments about students' reactions. In most cases, students were positive about the CBT experience.

- Excited to be in computer lab, even for testing
- More relaxed with CBT
- Prefer to have multiple-choice on computer
- Thought it was easy
- Really liked animations
- Very engaged

Other comments

- Biology test was too long.
- Some students had anxiety about using computer.
- Many thought it was difficult.

Additional comments regarding the DFA

Thirty-four additional comments were given about the *DFA*. Most found it clear and easy to use. Three asked for more detail in general, and the rest offered suggestions. Most commonly requested was more information about the Part 1/Part 2 transition. Also suggested by more than one respondent was to have a *DFA* for the tutorial and to provide more SAY box language for directing students how to use tickets to log on.

These *DFA* suggestions came individually:

- Remind students to use tools.
- Make sure the icon in the *DFA* matches the actual icon.
- More explanation on the length of the test
- Have a clear "NOW YOU MAY BEGIN" say box with nothing said after that.
- Add what to do if a question doesn't work or a screen freezes.
- Have a video about using *DFAs*.

Additional comments about student session tickets

Forty test administrators, or 47 percent, made additional comments about tickets. Twelve of those stated that someone else printed the tickets for them. Fourteen opted to print one ticket per page, and just one printed twelve per page. A couple of test administrators suggested adding an option for printing two tickets per page.

Four commented that the tickets were easy to use. Suggestions included:

- Print them larger so students don't lose them.
- Bold the user name and password.
- Put both sessions on one page.

One asked "Why are three codes needed for each test?"

A few test administrators mentioned it should be easier to group students by teacher or session. Some pointed out that their students did not need the tickets for Part 2 (in cases where students moved directly into Part 2 from Part 1). One mentioned problems with deleting the initial Pre-ID list and uploading a new list.

Additional comments

Twenty-one test administrators, or 25 percent, gave additional comments. The largest proportion gave positive feedback about their students' enthusiasm, ease of monitoring the CBT tests, not having to manage paper, and the convenience of instant scoring and results functionality. Quotes gathered included the following:

> *"It seems like this will work pretty easy. . . The amount of time this saves is amazing. I can see that the schools will be able to reduce personnel and staffing, and the tremendous amount of paper-pushing will be virtually non-existent."*

Several students were very enthusiastic. One quote that best sums up their perspective, "Bring it on!"

Eight administrators, or nine percent, expressed concern that there were not enough computers or capacity to roll CBT out to all students for all of STAR. Three mentioned the need to communicate better to students about what to do at the end of Part 1. Other concerns related to having enough time to set up the technology, scheduling, the need for a longer testing window, and the need for better organization and directions.

Other comments included the following:

- Students did not use the tools and resources at the top of the page
- Students did not use test-taking strategies

Responses from Technology Coordinators

Note: Technical issues are not confirmed and listed as reported in the survey results. For a summary of confirmed technical issues, see page 20.

Did you experience any problems with the STAR CBT?

Twenty-two technology coordinators provided additional open-ended comments.

- Pause/resume button did not work in some instances. (In most cases, students were able to resume testing from the same point following a system reboot.)
- One technology coordinator reported that students were sent back to the beginning of a test/section following a break, instead of allowing them to continue from where they left off.
- Three technology coordinators, or 14 percent, commented that they had problems connecting to the Internet or CA&L's system. Of these three, two experienced issues with wireless connections and laptops using batteries.

  - Using laptops and counting on wireless connectivity and batteries lasting proved a problematic setup.
  - A temporary disconnection from the network caused a student test to never reconnect. Student had to sign on with another password for Part 2 and lost the first part.
- Nine technology coordinators, or 41 percent, had comments related to the software installation required for the CBT program. The comments ranged from the installation being easy to install to being very time consuming and difficult to understand and install:
  - From a technical perspective the software ran great.
  - It worked well except for: (1) The RESUME option would freeze up and we would have to reboot the PC, (2) Automatic Windows Updates ran and wanted to reboot the PC, right before testing started.
  - Installing the system on every computer was inconvenient. I would prefer to have this program run totally web-based without the need for a host or server. This would allow schools more flexibility for students that bring their own device. This current system restricts testing to school equipment.
  - Was challenging to understand the goal of each software module. The instructions were not user friendly and required a lot of time from the Technology Department to set up. It took approximately eight hours to configure and set up one lab.
- Two technology coordinators, or 10 percent, commented that computers froze during testing and the proctors had difficulty rebooting and reconnecting to the exam:
  - A few computers froze on a simulation (moving graphic of mountains and the water cycle). We tried the escape key, etc. and nothing worked. We had to turn the computer off and on again and the student had to start over. [**Note:** This simulation came at the beginning of the test part, so no student data were lost.]
  - We had one student's test freeze up during testing. This was on a new Macbook laptop. We had to force quit to resolve the issue.
- One technology coordinator commented that without a mobile testing platform using iPads or Chromebooks, CBT testing would not be a practical format in their district.

Did you have any problems with the installation of the testing system or the Local Caching Software (LCS) system?

Sixteen technology coordinators, or 73 percent, had comments on the installation of the testing system or the LCS system. Many of the coordinators felt the system was very easy to install and had no problems activating the features, whereas some coordinators had difficulty installing the software or found the process too time consuming. Some technology coordinators felt the instructions for installing the program lacked detail.

- We had no problems with any part of the software installs.
- No problems at all. Very easy install.
- The installation was straightforward.
- No problems at all. Actually had student interns do the installation of the client, configuration, and testing of the connection to the LCS.
- Time consuming
- Could not run off of server so had to install on every computer.

- We really wanted to install the software remotely on each computer, but we just didn't have enough information or experience to get that approach to work.
- Inadequate instructions. Had to call tech support.

What technical support did you provide to schools during testing?

Twenty-five test administrators responded regarding the level of assistance they provided to schools during testing. Many LEAs had technical personnel ready to provide services to the schools; their involvement was minimal once testing started, as seen by the following comments:

- I conducted a test administration of the test for the teachers to ensure that the machines operated as intended. Once everything was in place the testing went very smoothly.
- District technicians were present at the time of the first assessment session.
- I sent technical support to schools to make sure everything was working. At one school the program kept timing out after 3 hours.

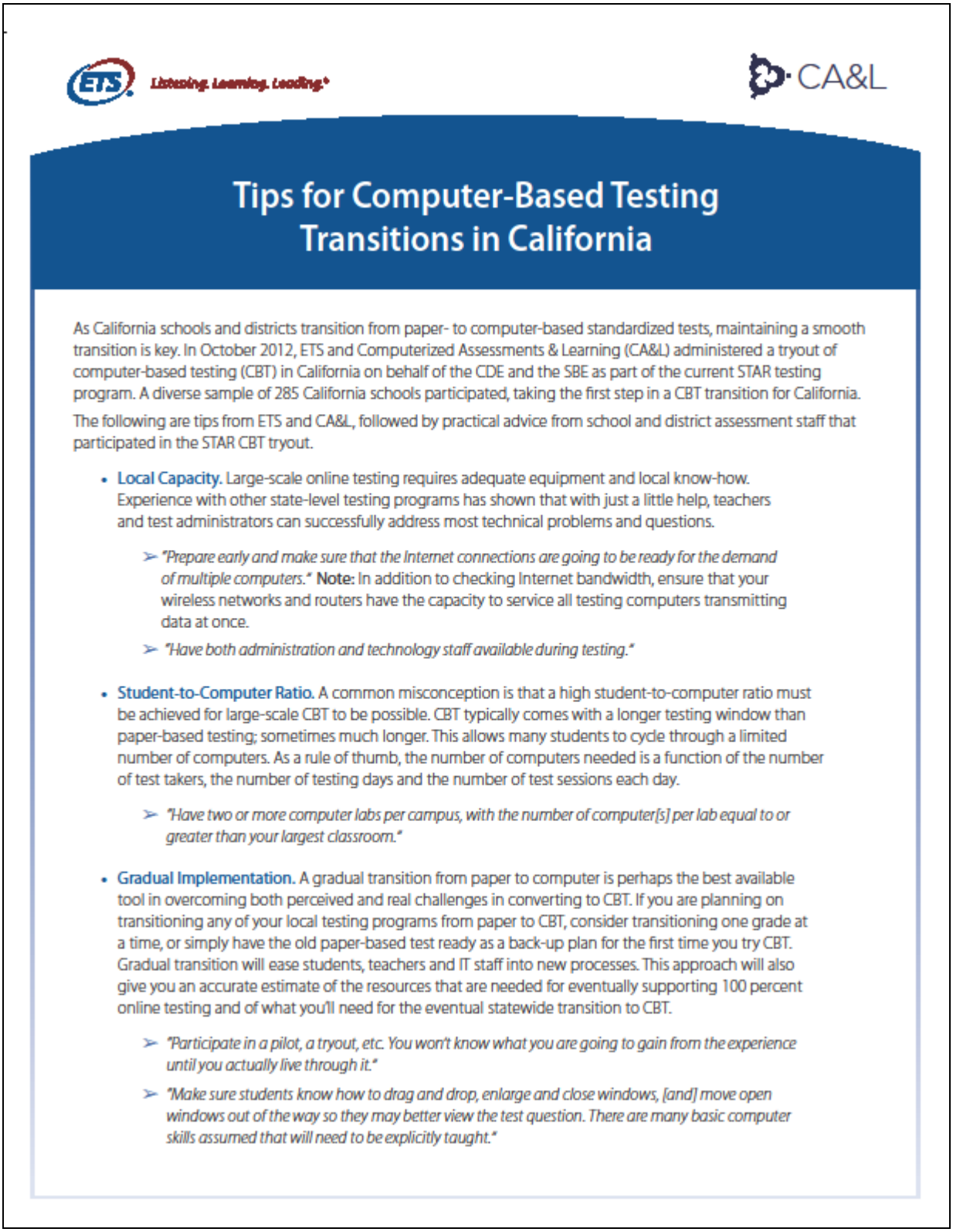
Do you have any additional comments or concerns regarding the technical aspect of the CBT Tryout?

Ten technical coordinators, or 45 percent, had a very broad range of suggestions and comments regarding the technical aspect of the CBT Tryout. What follow are some of their responses:

- (1) Would be nice to track students' progress throughout the test instead of just seeing start and stop. (2) Can see that computer availability and the choreography of student testing will become more challenging as the number of students testing becomes larger.
- Without a mobile platform (iPads, Chromebooks, Android) this system cannot be considered as viable for deployment.
- The technical support staff that assisted were very helpful. In a couple of minutes they had me back on track.
- Everything went as planned – the server/client model worked very well for this application. Things went very smooth for us from deployment to test administration.

Appendix D—Documentation

Tip Sheet



Listening. Learning. Leading.®

CA&L

Tips for Computer-Based Testing Transitions in California

As California schools and districts transition from paper- to computer-based standardized tests, maintaining a smooth transition is key. In October 2012, ETS and Computerized Assessments & Learning (CA&L) administered a tryout of computer-based testing (CBT) in California on behalf of the CDE and the SBE as part of the current STAR testing program. A diverse sample of 285 California schools participated, taking the first step in a CBT transition for California.

The following are tips from ETS and CA&L, followed by practical advice from school and district assessment staff that participated in the STAR CBT tryout.

- **Local Capacity.** Large-scale online testing requires adequate equipment and local know-how. Experience with other state-level testing programs has shown that with just a little help, teachers and test administrators can successfully address most technical problems and questions.
  - ➢ *"Prepare early and make sure that the Internet connections are going to be ready for the demand of multiple computers."* **Note:** In addition to checking Internet bandwidth, ensure that your wireless networks and routers have the capacity to service all testing computers transmitting data at once.
  - ➢ *"Have both administration and technology staff available during testing."*
- **Student-to-Computer Ratio.** A common misconception is that a high student-to-computer ratio must be achieved for large-scale CBT to be possible. CBT typically comes with a longer testing window than paper-based testing; sometimes much longer. This allows many students to cycle through a limited number of computers. As a rule of thumb, the number of computers needed is a function of the number of test takers, the number of testing days and the number of test sessions each day.
  - ➢ *"Have two or more computer labs per campus, with the number of computer[s] per lab equal to or greater than your largest classroom."*
- **Gradual Implementation.** A gradual transition from paper to computer is perhaps the best available tool in overcoming both perceived and real challenges in converting to CBT. If you are planning on transitioning any of your local testing programs from paper to CBT, consider transitioning one grade at a time, or simply have the old paper-based test ready as a back-up plan for the first time you try CBT. Gradual transition will ease students, teachers and IT staff into new processes. This approach will also give you an accurate estimate of the resources that are needed for eventually supporting 100 percent online testing and of what you'll need for the eventual statewide transition to CBT.
  - ➢ *"Participate in a pilot, a tryout, etc. You won't know what you are going to gain from the experience until you actually live through it."*
  - ➢ *"Make sure students know how to drag and drop, enlarge and close windows, [and] move open windows out of the way so they may better view the test question. There are many basic computer skills assumed that will need to be explicitly taught."*

- **Training and Communication.** Teachers, administrators and technology staff each have their own roles in the successful administration of computer-based assessments and adequate training and communication of these roles will be key. Students will also need training.
  - ➢ *"Do not stress or panic. My advice would be to stay updated on all correspondence from [the Smarter Balanced Assessment Consortium], CDE and STAR. Share all information with your administrators and teachers."*
  - ➢ *"It helps if the IT department and Education Services department have a good working relationship. Communication and cooperation were key for us."*
  - ➢ *"Be aware that while most students were more than comfortable with the CBT setting, the students that do not have computers at home became more apparent. The tutorial was a 'must watch'."*
- **Involve Technology Staff Early and Often.** Early participation by IT staff ensures computers and systems can be thoroughly prepared and checked before test day. Each district network is unique, and custom setup procedures may be necessary.
  - ➢ *"Work early with the district IT department. The sooner they load the software, the more time the teachers and students will have to do the practice tests."*
  - ➢ *"Involve your tech people early!"*
  - ➢ *"The technology readiness piece is incredible. We would not have been as successful had we not had total tech support at each of the testing sessions."*
  - ➢ *"Technology support personnel need to be an integral part of the process."*
- **Discuss the Advantages of Computer-Based Testing.** There are many advantages to CBT beyond the simple elimination of large-scale paper handling. Discuss these advantages with your staff. They include:
  - ✓ Potential for immediate test results
  - ✓ Accuracy in data and data collection
  - ✓ Student motivation
  - ✓ Use of new item types that better assess student knowledge and skills
  - ✓ Leverage for added hardware purchases that will also support instruction
  - ✓ Pre-ID flexibility, which allows test-taker information to be added and changed on the fly
  - ✓ Fairer assessments through technology-enabled modifications and accommodations
  - ➢ *"[You] need to try it. Our students would much prefer to take [tests] online versus paper. They said it was easier to navigate, so this is the wave of their future."*
  - ➢ *"Give it a try!"*

ETS

Listening. Learning. Leading.®

www.ets.org

Appendix E—Student Response Data and Aggregate Score Results

CBT Item Statistics

Table E.1. CBT Item Statistics for Grade Five Science—CBT vs. Previous PPT Administrations

	N Students		*p*-value		Point-Biserial		Pct Omitted	
CBT Item Number	PPT	CBT †	PPT	CBT	PPT	CBT	PPT	CBT
1	364,436	9,656	0.68	0.79	0.47	0.37	0.17	0.14
2	191,012	9,656	0.51	0.52	0.33	0.18	0.16	0.22
3 *	N/A	9,655	N/A	0.04	N/A	0.10	N/A	0.49
4	203,500	9,653	0.38	0.41	0.26	0.27	0.20	0.36
5	203,500	9,653	0.36	0.33	0.39	0.39	0.12	0.31
6	191,012	9,653	0.31	0.41	0.33	0.41	0.19	0.30
7	203,500	9,653	0.33	0.40	0.35	0.40	0.22	0.40
8	364,436	9,653	0.67	0.72	0.39	0.33	0.14	0.44
9	203,500	9,652	0.51	0.51	0.29	0.27	0.15	0.38
10 *	N/A	9,650	N/A	0.03	N/A	0.22	N/A	0.50
11	21,811	9,650	0.33	0.31	0.25	0.19	0.17	0.39
12	397,570	9,650	0.60	0.55	0.43	0.47	0.12	0.54
13	191,012	9,650	0.25	0.34	0.23	0.34	0.07	0.44
14 *	N/A	9,648	N/A	0.14	N/A	0.25	N/A	1.45
15	191,012	9,648	0.34	0.45	0.28	0.29	0.18	0.59
16 *	N/A	9,648	N/A	0.67	N/A	0.41	N/A	0.53
17 *	N/A	9,647	N/A	0.92	N/A	0.21	N/A	0.64
18 *	N/A	9,647	N/A	0.94	N/A	0.17	N/A	0.64
19	191,012	9,645	0.47	0.49	0.35	0.39	0.12	0.94
20	397,570	9,643	0.73	0.80	0.48	0.37	0.07	0.93
21	191,012	9,643	0.66	0.72	0.52	0.47	0.12	1.09
22 *	N/A	9,635	N/A	0.13	N/A	0.32	N/A	1.35
23	191,012	9,634	0.42	0.53	0.28	0.38	0.18	1.70
24 *	N/A	9,634	N/A	0.15	N/A	0.24	N/A	1.89
25	203,500	9,633	0.59	0.54	0.53	0.51	0.17	1.82
26	191,012	9,629	0.51	0.48	0.41	0.39	0.37	2.01
27 *	N/A	9,628	N/A	0.58	N/A	0.30	N/A	2.03
28	203,500	9,626	0.45	0.48	0.30	0.28	0.08	2.26
29	21,811	9,623	0.55	0.46	0.40	0.40	0.17	2.39
30 *	N/A	9,614	N/A	0.06	N/A	0.19	N/A	2.38
31 *	N/A	9,084	N/A	0.01	N/A	0.04	N/A	0.68
32 *	N/A	9,084	N/A	0.42	N/A	0.41	N/A	0.75
33 *	N/A	9,084	N/A	0.80	N/A	0.42	N/A	0.76
34 *	N/A	9,084	N/A	0.45	N/A	0.28	N/A	0.75
35 *	N/A	9,084	N/A	0.50	N/A	0.45	N/A	0.80
36	378,641	9,051	0.79	0.84	0.43	0.36	0.13	0.76
37	378,641	9,051	0.77	0.79	0.38	0.30	0.18	0.82
38 *	N/A	9,050	N/A	0.33	N/A	0.23	N/A	0.84
39	203,500	9,050	0.40	0.49	0.22	0.24	0.17	0.86
40	386,351	9,049	0.65	0.60	0.40	0.40	0.12	0.84
41	364,436	9,047	0.72	0.77	0.48	0.46	0.07	0.85

CBT Item Number	N Students		*p*-value		Point-Biserial		Pct Omitted	
	PPT	CBT †	PPT	CBT	PPT	CBT	PPT	CBT
42	203,500	9,044	0.63	0.60	0.39	0.37	0.10	0.92
43	191,012	9,043	0.55	0.63	0.45	0.50	0.17	0.91
44	364,436	9,041	0.63	0.64	0.38	0.39	0.17	0.96
45 *	N/A	9,040	N/A	0.44	N/A	0.26	N/A	11.78
46	203,500	9,036	0.49	0.43	0.25	0.26	0.17	1.18
47	378,641	9,036	0.65	0.61	0.45	0.43	0.11	1.14
48	203,500	9,035	0.52	0.50	0.37	0.37	0.13	1.26
49 *	N/A	9,035	N/A	0.47	N/A	0.53	N/A	1.24
50	203,500	9,035	0.50	0.43	0.40	0.42	0.14	1.24
51	191,012	9,034	0.48	0.50	0.39	0.44	0.12	1.31
52 *	N/A	9,031	N/A	0.25	N/A	0.37	N/A	1.57
53	397,570	9,031	0.82	0.84	0.36	0.38	0.06	1.43
54 *	N/A	9,030	N/A	0.64	N/A	0.27	N/A	3.06
55	191,012	9,024	0.51	0.51	0.36	0.41	0.15	1.85
56	203,500	9,024	0.71	0.67	0.47	0.43	0.10	1.93
57	191,012	9,024	0.53	0.60	0.33	0.35	0.06	1.94
58	191,012	9,022	0.55	0.59	0.36	0.44	0.17	2.07
59	364,436	9,019	0.60	0.51	0.38	0.37	0.07	2.04
60	191,012	9,018	0.60	0.64	0.36	0.39	0.17	2.05
Reused Item Correlation	N/A		0.92		0.74		–0.04	

* Technology-enhanced item

† N listed is for CBT p-value and percent omitted (**Pct Omitted**), which is the number of students tested (9,659) minus the number of not-presented cases for that item. The number of students used for point-biserials is 9,018, which is the number of students with a valid total score (i.e., none of the item responses were marked as "not presented").

Table E.2. CBT Item Statistics for Grade Eight Science—CBT vs. Previous PPT Administrations

CBT Item Number	N Students		*p*-value		Point-Biserial		Pct Omitted	
	PPT	CBT †	PPT	CBT	PPT	CBT	PPT	CBT
1	397,427	6,031	0.72	0.73	0.38	0.33	0.04	0.32
2	382,648	6,031	0.73	0.79	0.35	0.27	0.07	0.46
3 *	N/A	6,031	N/A	0.71	N/A	0.47	N/A	0.90
4	219,032	6,031	0.30	0.27	0.32	0.36	0.17	0.58
5	218,032	6,030	0.50	0.52	0.34	0.30	0.14	0.66
6	382,648	6,029	0.73	0.75	0.43	0.37	0.11	0.65
7	218,032	6,028	0.65	0.86	0.54	0.43	0.15	0.55
8 *	N/A	6,027	N/A	0.21	N/A	0.48	N/A	0.68
9	218,032	6,025	0.35	0.35	0.29	0.29	0.16	0.66
10	218,032	6,024	0.44	0.49	0.31	0.29	0.09	0.90
11 *	N/A	6,024	N/A	0.70	N/A	0.46	N/A	0.75
12	218,032	6,023	0.44	0.46	0.44	0.49	0.18	0.93
13	382,648	6,023	0.61	0.54	0.38	0.27	0.13	0.91
14	219,032	6,022	0.38	0.38	0.19	0.22	0.16	1.06
15	219,032	6,022	0.46	0.52	0.42	0.45	0.06	1.03
16 *	N/A	6,021	N/A	0.56	N/A	0.49	N/A	1.08
17	219,032	6,019	0.41	0.47	0.32	0.39	0.15	1.43
18	219,032	6,019	0.35	0.38	0.22	0.26	0.21	1.68
19 *	N/A	6,018	N/A	0.64	N/A	0.46	N/A	1.66

	N Students		*p*-value		Point-Biserial		Pct Omitted	
CBT Item Number	PPT	CBT †	PPT	CBT	PPT	CBT	PPT	CBT
20	218,032	6,018	0.72	0.87	0.30	0.26	0.09	1.71
21	382,648	6,016	0.72	0.75	0.43	0.37	0.09	1.81
22 *	N/A	6,015	N/A	0.42	N/A	0.42	N/A	1.80
23	397,427	6,015	0.53	0.57	0.42	0.36	0.08	1.81
24 *	N/A	6,012	N/A	0.35	N/A	0.42	N/A	4.56
25	219,032	6,012	0.26	0.36	0.23	0.33	0.19	2.35
26 *	N/A	6,012	N/A	0.14	N/A	0.50	N/A	3.93
27	219,032	6,012	0.35	0.44	0.32	0.43	0.27	2.89
28 *	N/A	6,011	N/A	0.17	N/A	0.45	N/A	3.96
29	219,032	6,011	0.59	0.67	0.48	0.46	0.16	3.19
30 *	N/A	6,011	N/A	0.71	N/A	0.45	N/A	3.39
31 *	N/A	5,713	N/A	0.57	N/A	0.54	N/A	1.14
32 *	N/A	5,713	N/A	0.20	N/A	0.51	N/A	1.19
33 *	N/A	5,713	N/A	0.72	N/A	0.40	N/A	1.21
34 *	N/A	5,713	N/A	0.03	N/A	0.18	N/A	1.21
35 *	N/A	5,713	N/A	0.25	N/A	0.55	N/A	1.38
36	219,032	5,701	0.62	0.78	0.48	0.43	0.13	1.60
37	397,427	5,701	0.75	0.84	0.31	0.30	0.07	1.58
38	219,032	5,699	0.39	0.42	0.24	0.26	0.25	1.70
39 *	N/A	5,690	N/A	0.42	N/A	0.42	N/A	1.95
40 *	N/A	5,688	N/A	0.33	N/A	0.29	N/A	1.99
41	218,032	5,688	0.71	0.59	0.50	0.46	0.13	1.93
42	405,093	5,685	0.87	0.84	0.51	0.47	0.07	1.93
43	219,032	5,683	0.68	0.73	0.39	0.50	0.15	1.99
44	218,032	5,681	0.64	0.67	0.50	0.51	0.09	2.11
45	218,032	5,677	0.40	0.38	0.29	0.30	0.24	2.18
46	397,427	5,677	0.64	0.68	0.38	0.44	0.11	2.18
47	397,427	5,675	0.55	0.54	0.39	0.37	0.16	2.41
48 *	N/A	5,671	N/A	0.38	N/A	0.48	N/A	2.89
49 *	N/A	5,662	N/A	0.24	N/A	0.43	N/A	2.86
50	405,093	5,661	0.89	0.88	0.43	0.37	0.07	2.84
51	218,032	5,659	0.31	0.33	0.40	0.46	0.09	2.83
52	218,032	5,658	0.52	0.59	0.42	0.49	0.16	2.88
53	397,427	5,657	0.66	0.68	0.47	0.45	0.06	2.85
54 *	N/A	5,654	N/A	0.67	N/A	0.60	N/A	3.08
55	397,427	5,653	0.48	0.50	0.35	0.43	0.13	3.02
56	219,032	5,652	0.40	0.52	0.38	0.49	0.19	3.11
57	219,032	5,650	0.41	0.55	0.34	0.51	0.13	3.15
58	388,344	5,650	0.75	0.71	0.47	0.46	0.10	3.26
59	397,427	5,648	0.55	0.46	0.30	0.37	0.15	3.19
60	405,093	5,648	0.72	0.76	0.41	0.42	0.09	3.12
Reused Item Correlation	N/A		0.92		0.70		0.12	

* Technology-enhanced item

† N listed is for CBT p-value and percent omitted (**Pct Omitted**), which is the number of students tested (6,032) minus the number of not-presented cases for that item. The number of students used for point-biserials is 5,648, which is the number of students with a valid total score (i.e., none of the item responses were marked as "not presented").

Table E.3. CBT Item Statistics for Biology—CBT vs. Previous PPT Administrations

	N Students		*p*-value		Point-Biserial		Pct Omitted	
CBT Item Number	PPT	CBT †	PPT	CBT	PPT	CBT	PPT	CBT
1	245,799	5,389	0.82	0.89	0.48	0.33	0.07	0.09
2	245,799	5,389	0.45	0.39	0.30	0.32	0.18	0.22
3 *	N/A	5,389	N/A	0.12	N/A	0.09	N/A	0.30
4	245,799	5,387	0.58	0.66	0.43	0.41	0.15	0.19
5	518,271	5,387	0.51	0.45	0.31	0.26	0.22	0.26
6	294,783	5,387	0.53	0.61	0.45	0.42	0.11	0.24
7	456,724	5,387	0.61	0.63	0.36	0.40	0.12	0.24
8 *	N/A	5,387	N/A	0.16	N/A	0.40	N/A	0.43
9	245,799	5,387	0.40	0.48	0.33	0.28	0.21	0.32
10	294,783	5,387	0.51	0.48	0.35	0.26	0.11	0.30
11 *	N/A	5,386	N/A	0.38	N/A	0.33	N/A	0.33
12	294,783	5,386	0.51	0.49	0.40	0.37	0.08	0.26
13 *	N/A	5,386	N/A	0.55	N/A	0.35	N/A	0.33
14	245,799	5,386	0.58	0.57	0.39	0.30	0.10	0.28
15 *	N/A	5,386	N/A	0.54	N/A	0.39	N/A	0.26
16	245,799	5,386	0.41	0.41	0.28	0.22	0.23	0.37
17	245,799	5,386	0.63	0.70	0.51	0.48	0.10	0.32
18	245,799	5,386	0.50	0.48	0.42	0.38	0.29	0.37
19	518,271	5,386	0.61	0.54	0.47	0.43	0.13	0.46
20 *	N/A	5,386	N/A	0.21	N/A	0.38	N/A	0.59
21	294,783	5,386	0.56	0.52	0.51	0.28	0.14	0.46
22 *	N/A	5,385	N/A	0.38	N/A	0.46	N/A	0.91
23	245,799	5,384	0.36	0.32	0.27	0.31	0.17	0.54
24	245,799	5,384	0.72	0.73	0.54	0.43	0.17	0.56
25	456,724	5,384	0.57	0.60	0.45	0.41	0.13	0.67
26 *	N/A	5,384	N/A	0.36	N/A	0.47	N/A	0.71
27 *	N/A	5,384	N/A	0.60	N/A	0.35	N/A	0.69
28	456,724	5,383	0.47	0.49	0.41	0.41	0.19	0.71
29	456,724	5,383	0.72	0.71	0.46	0.35	0.09	0.76
30	294,783	5,383	0.58	0.47	0.33	0.35	0.14	0.74
31	456,724	5,383	0.59	0.50	0.35	0.35	0.06	0.74
32 *	N/A	5,145	N/A	0.49	N/A	0.53	N/A	1.75
33 *	N/A	5,145	N/A	0.20	N/A	0.46	N/A	1.50
34 *	N/A	5,145	N/A	0.23	N/A	0.30	N/A	1.50
35 *	N/A	5,145	N/A	0.43	N/A	0.56	N/A	1.77
36 *	N/A	5,145	N/A	0.21	N/A	0.49	N/A	2.08
37 *	N/A	5,145	N/A	0.56	N/A	0.62	N/A	2.27
38 *	N/A	5,145	N/A	0.17	N/A	0.15	N/A	2.41
39	294,783	5,103	0.53	0.44	0.50	0.45	0.16	1.16
40	294,783	5,100	0.51	0.53	0.50	0.53	0.14	1.18
41 *	N/A	5,098	N/A	0.54	N/A	0.26	N/A	1.16
42	455,116	5,091	0.63	0.63	0.53	0.49	0.13	1.16
43	245,799	5,084	0.61	0.68	0.48	0.45	0.21	1.20
44 *	N/A	5,084	N/A	0.11	N/A	0.27	N/A	1.30
45	245,799	5,083	0.50	0.49	0.39	0.32	0.18	1.32
46	245,799	5,081	0.56	0.60	0.54	0.50	0.07	1.30
47	393,361	5,080	0.70	0.68	0.34	0.25	0.13	1.38
48 *	N/A	5,070	N/A	0.42	N/A	0.55	N/A	2.29

CBT Item Number	N Students		*p*-value		Point-Biserial		Pct Omitted	
	PPT	CBT †	PPT	CBT	PPT	CBT	PPT	CBT
49	294,783	5,067	0.34	0.28	0.35	0.34	0.10	1.58
50	456,724	5,067	0.84	0.81	0.39	0.45	0.08	1.60
51	518,271	5,066	0.62	0.55	0.31	0.29	0.09	1.60
52	294,783	5,065	0.56	0.59	0.49	0.47	0.12	1.64
53	456,724	5,063	0.67	0.66	0.53	0.49	0.15	1.62
54 *	N/A	5,061	N/A	0.46	N/A	0.41	N/A	3.93
55	245,799	5,061	0.41	0.36	0.30	0.39	0.24	1.90
56	245,799	5,060	0.57	0.55	0.40	0.40	0.09	1.86
57	245,799	5,059	0.56	0.47	0.54	0.46	0.16	1.88
58	294,783	5,057	0.54	0.48	0.50	0.49	0.18	1.96
59	245,799	5,056	0.59	0.58	0.45	0.40	0.17	1.96
60	456,724	5,051	0.53	0.50	0.38	0.34	0.17	1.98
Reused Item Correlation	N/A		0.92		0.75		0.01	

* Technology-enhanced item

† N listed is for CBT *p*-value and percent omitted (**Pct Omitted**), which is the number of students tested (5,404) minus the number of not-presented cases for that item. The number of students used for point-biserials is 5,051, which is the number of students with a valid total score (i.e. none of the item responses were marked as "not presented").

Percent Omit by Subgroups

Table E.4. CBT Percent Omit for Grade Five by EL Status and Economic Status

CBT Item Number	English Learner	Non–English Learner	Economically Disadvantaged	Not Economically Disadvantaged
1	0.26	0.09	0.13	0.12
2	0.30	0.16	0.20	0.17
3 *	0.47	0.44	0.40	0.52
4	0.38	0.31	0.32	0.35
5	0.47	0.21	0.32	0.20
6	0.47	0.20	0.29	0.23
7	0.55	0.31	0.40	0.32
8	0.68	0.31	0.35	0.50
9	0.55	0.27	0.32	0.41
10 *	0.68	0.40	0.47	0.47
11	0.60	0.29	0.34	0.41
12	0.77	0.43	0.44	0.64
13	0.64	0.33	0.37	0.47
14 *	1.45	1.37	1.23	1.66
15	0.81	0.49	0.57	0.55
16 *	0.72	0.44	0.52	0.50
17 *	0.85	0.53	0.64	0.55
18 *	0.85	0.53	0.64	0.55
19	1.11	0.83	0.93	0.85
20	1.15	0.80	0.96	0.76
21	1.53	0.89	1.10	0.96
22 *	1.62	1.21	1.33	1.29
23	2.13	1.49	1.70	1.55
24*	2.56	1.59	1.96	1.61
25	2.39	1.56	1.92	1.49
26	2.56	1.77	2.08	1.78
27 *	2.77	1.72	2.09	1.78
28	3.02	1.96	2.30	2.11
29	3.11	2.10	2.48	2.14
30 *	3.07	2.09	2.43	2.17
31 *	0.40	0.74	0.51	0.92
32 *	0.45	0.82	0.55	1.04
33 *	0.49	0.82	0.58	1.01
34 *	0.40	0.84	0.57	1.01
35 *	0.58	0.85	0.64	1.04
36	0.67	0.75	0.57	1.02
37	0.72	0.81	0.60	1.11
38 *	0.76	0.81	0.62	1.11
39	0.85	0.82	0.69	1.08
40	0.76	0.82	0.64	1.11
41	0.81	0.82	0.66	1.11
42	0.85	0.90	0.71	1.21
43	0.90	0.87	0.73	1.14
44	0.95	0.92	0.78	1.18
45 *	15.53	10.40	12.51	10.24
46	1.22	1.12	1.03	1.34

CBT Item Number	English Learner	Non–English Learner	Economically Disadvantaged	Not Economically Disadvantaged
47	1.31	1.01	1.00	1.24
48	1.44	1.13	1.16	1.31
49 *	1.49	1.08	1.12	1.31
50	1.44	1.12	1.16	1.27
51	1.58	1.16	1.19	1.40
52 *	1.76	1.47	1.42	1.75
53	1.67	1.28	1.35	1.43
54 *	3.33	2.81	2.99	2.87
55	1.98	1.74	1.87	1.69
56	2.03	1.84	1.92	1.82
57	2.03	1.85	1.96	1.78
58	2.16	1.96	2.08	1.88
59	2.25	1.91	2.01	1.98
60	2.30	1.91	2.05	1.94

* Technology-enhanced item

Table E.5. CBT Percent Omit for Grade Eight by EL Status and Economic Status

CBT Item Number	English Learner	Non–English Learner	Economically Disadvantaged	Not Economically Disadvantaged
1	0.13	0.09	0.09	0.10
2	0.39	0.22	0.21	0.30
3 *	1.30	0.50	0.69	0.50
4	0.52	0.31	0.33	0.35
5	0.65	0.39	0.45	0.40
6	0.65	0.42	0.51	0.35
7	0.65	0.28	0.36	0.30
8 *	0.78	0.39	0.51	0.35
9	0.78	0.39	0.45	0.45
10	1.04	0.63	0.78	0.55
11 *	1.04	0.46	0.63	0.40
12	1.30	0.63	0.75	0.69
13	1.43	0.59	0.81	0.55
14	1.56	0.79	1.05	0.65
15	1.69	0.70	0.96	0.65
16 *	2.08	0.70	1.05	0.65
17	2.60	1.03	1.41	0.99
18	2.34	1.33	1.74	1.04
19 *	2.60	1.27	1.65	1.14
20	2.60	1.31	1.68	1.19
21	2.61	1.44	1.83	1.24
22 *	2.74	1.42	1.83	1.24
23	2.87	1.40	1.80	1.29
24 *	4.83	4.13	4.48	3.82
25	3.66	1.95	2.59	1.54
26 *	4.44	3.39	3.94	2.88
27	4.31	2.41	3.07	2.04
28 *	6.01	3.26	4.12	2.88
29	4.96	2.63	3.46	2.14
30 *	5.09	2.78	3.64	2.24

CBT Item Number	English Learner	Non–English Learner	Economically Disadvantaged	Not Economically Disadvantaged
31 *	1.25	0.96	1.01	0.98
32 *	1.39	1.00	1.07	1.03
33 *	1.52	1.00	1.10	1.03
34 *	1.52	1.00	1.10	1.03
35 *	2.08	1.14	1.29	1.24
36	2.51	1.19	1.55	1.08
37	2.51	1.14	1.52	1.03
38	2.66	1.25	1.58	1.24
39 *	3.23	1.39	1.87	1.29
40	2.81	1.44	1.74	1.45
41	2.81	1.39	1.68	1.45
42	2.67	1.39	1.65	1.45
43	2.81	1.44	1.68	1.55
44	2.67	1.53	1.81	1.50
45	2.81	1.58	1.87	1.55
46	2.95	1.51	1.81	1.55
47	3.09	1.63	1.94	1.66
48 *	3.66	2.04	2.26	2.28
49 *	3.80	1.79	2.13	1.98
50	3.66	1.70	2.07	1.82
51	3.66	1.68	2.07	1.77
52	3.94	1.70	2.13	1.82
53	3.80	1.68	2.10	1.77
54 *	3.94	1.93	2.39	1.93
55	3.80	1.79	2.20	1.87
56	3.80	1.89	2.26	1.98
57	3.80	1.89	2.33	1.87
58	3.80	1.95	2.42	1.87
59	3.81	1.86	2.30	1.87
60	3.67	1.84	2.33	1.72

* Technology-enhanced item

Table E.6. CBT Percent Omit for Biology by EL Status and Economic Status

CBT Item Number	English Learner	Non–English Learner	Economically Disadvantaged	Not Economically Disadvantaged
1	0.18	0.09	0.13	0.05
2	0.18	0.23	0.26	0.18
3 *	0.36	0.28	0.29	0.27
4	0.18	0.19	0.20	0.18
5	0.18	0.28	0.29	0.23
6	0.18	0.23	0.26	0.18
7	0.18	0.23	0.23	0.23
8 *	0.18	0.45	0.43	0.41
9	0.18	0.32	0.33	0.27
10	0.18	0.30	0.33	0.23
11 *	0.18	0.34	0.33	0.32
12	0.18	0.26	0.26	0.23
13	0.18	0.32	0.33	0.27
14	0.18	0.28	0.29	0.23

CBT Item Number	English Learner	Non–English Learner	Economically Disadvantaged	Not Economically Disadvantaged
15 *	0.36	0.23	0.29	0.18
16	0.18	0.38	0.39	0.32
17	0.18	0.32	0.33	0.27
18	0.36	0.36	0.46	0.23
19	0.36	0.47	0.49	0.41
20 *	0.54	0.60	0.72	0.41
21	0.54	0.45	0.52	0.36
22 *	1.27	0.85	1.02	0.73
23	0.72	0.51	0.66	0.36
24	0.54	0.55	0.66	0.41
25	0.54	0.64	0.72	0.50
26 *	0.54	0.70	0.75	0.59
27 *	0.54	0.68	0.79	0.50
28	0.72	0.68	0.82	0.50
29	0.90	0.72	0.95	0.45
30	0.72	0.72	0.85	0.54
31	0.72	0.72	0.89	0.50
32 *	3.20	1.56	1.85	1.57
33 *	1.69	1.47	1.54	1.43
34 *	1.69	1.47	1.54	1.43
35 *	2.26	1.72	1.81	1.72
36 *	2.44	2.05	2.02	2.19
37 *	2.82	2.23	2.29	2.29
38 *	2.82	2.38	2.43	2.43
39	0.95	1.17	1.04	1.30
40	0.95	1.19	0.93	1.49
41 *	0.95	1.17	0.93	1.44
42	0.95	1.17	0.93	1.44
43	1.52	1.15	0.97	1.50
44 *	1.14	1.29	1.07	1.54
45	1.14	1.31	1.11	1.55
46	1.14	1.29	1.07	1.55
47	1.14	1.35	1.07	1.69
48 *	2.48	2.22	2.12	2.42
49	1.34	1.56	1.32	1.84
50	1.34	1.58	1.32	1.89
51	1.34	1.58	1.36	1.84
52	1.34	1.63	1.39	1.89
53	1.34	1.61	1.39	1.84
54 *	4.99	3.78	3.90	3.92
55	1.54	1.90	1.64	2.18
56	1.34	1.88	1.57	2.18
57	1.34	1.90	1.60	2.18
58	1.34	1.99	1.74	2.18
59	1.34	1.97	1.67	2.23
60	1.35	2.00	1.71	2.23

* Technology-enhanced item

Percent Omit by Proficiency on the Paper-and-Pencil Tests

Table E.7. CBT Percent Omit for Grade Five by Performance Level on the Spring 2012 CST for TE and MC Items

CBT Item Number	Far Below Basic	Below Basic	Basic	Proficient	Advanced
1	0.00	0.13	0.18	0.16	0.00
2	0.40	0.13	0.23	0.22	0.09
3 *	0.40	0.53	0.45	0.62	0.17
4	0.00	0.27	0.32	0.47	0.21
5	0.00	0.27	0.32	0.41	0.09
6	0.00	0.27	0.32	0.34	0.13
7	0.00	0.40	0.50	0.44	0.13
8	0.00	0.53	0.41	0.56	0.17
9	0.00	0.27	0.45	0.41	0.21
10 *	0.20	0.67	0.45	0.56	0.30
11	0.00	0.27	0.41	0.47	0.21
12	0.20	0.27	0.50	0.72	0.34
13	0.00	0.27	0.41	0.56	0.26
14 *	1.01	0.93	1.31	1.50	1.41
15	0.00	0.53	0.59	0.75	0.34
16 *	0.20	0.67	0.45	0.66	0.30
17 *	0.20	0.80	0.72	0.66	0.30
18 *	0.20	0.80	0.72	0.66	0.30
19	0.20	1.20	1.13	0.91	0.51
20	0.20	1.20	1.13	0.84	0.51
21	0.40	1.60	1.40	0.97	0.56
22 *	0.61	1.47	1.67	1.31	0.81
23	0.40	1.74	1.81	1.84	1.15
24*	1.41	1.74	2.04	2.00	1.07
25	0.81	1.74	1.99	1.94	1.11
26	1.41	1.87	2.08	2.22	1.28
27 *	1.21	1.74	2.04	2.28	1.41
28	1.21	2.14	2.36	2.53	1.67
29	1.41	2.27	2.54	2.63	1.71
30 *	1.41	2.40	2.40	2.70	1.63
31 *	1.06	0.28	0.48	0.70	0.68
32 *	1.06	0.28	0.53	0.80	0.73
33 *	1.48	0.28	0.53	0.73	0.82
34 *	1.06	0.28	0.53	0.76	0.82
35 *	1.27	0.71	0.57	0.76	0.82
36	0.42	0.85	0.43	0.90	0.64
37	0.42	1.00	0.48	1.00	0.64
38 *	1.06	0.85	0.48	0.93	0.68
39	0.64	0.85	0.53	1.00	0.78
40	0.64	0.85	0.43	1.03	0.73
41	0.85	0.85	0.48	0.97	0.78
42	0.85	0.85	0.57	1.10	0.78
43	0.85	0.85	0.53	1.07	0.82
44	0.85	0.85	0.58	1.10	0.91

CBT Item Number	Far Below Basic	Below Basic	Basic	Proficient	Advanced
45 *	21.44	18.52	14.00	10.14	5.85
46	1.49	1.00	0.82	1.27	1.14
47	1.49	1.14	0.86	1.10	1.05
48	1.70	1.14	0.96	1.30	1.10
49 *	1.49	1.57	1.01	1.20	1.01
50	1.70	1.28	0.91	1.27	1.10
51	1.91	1.42	0.91	1.37	1.14
52 *	1.91	1.85	1.15	1.70	1.37
53	2.12	1.57	0.91	1.60	1.14
54 *	4.26	4.27	2.98	2.71	2.24
55	2.35	1.57	1.49	1.97	1.78
56	2.56	1.57	1.54	2.01	1.92
57	2.35	1.57	1.49	2.11	1.97
58	2.56	1.85	1.63	2.21	2.01
59	2.35	1.85	1.63	2.21	2.06
60	2.14	1.71	1.82	2.17	2.06

* Technology-enhanced item

Table E.8. CBT Percent Omit for Grade Eight by Performance Level on the Spring 2012 CST for TE and MC Items

CBT Item Number	Far Below Basic	Below Basic	Basic	Proficient	Advanced
1	0.30	0.00	0.00	0.00	0.12
2	0.61	0.51	0.13	0.00	0.27
3 *	2.44	1.01	0.40	0.37	0.31
4	0.91	0.25	0.00	0.37	0.35
5	0.61	0.51	0.13	0.28	0.46
6	0.61	0.76	0.27	0.37	0.39
7	0.91	0.25	0.00	0.28	0.31
8 *	1.52	0.25	0.13	0.47	0.31
9	0.91	0.51	0.13	0.37	0.42
10	1.22	0.76	0.40	0.65	0.58
11 *	1.22	0.76	0.40	0.47	0.39
12	1.22	0.51	0.40	0.47	0.77
13	0.91	0.51	0.40	0.75	0.66
14	1.22	1.27	0.53	0.94	0.73
15	1.22	0.51	0.67	0.65	0.81
16 *	0.91	1.52	0.93	0.75	0.69
17	1.83	1.27	1.47	1.40	0.93
18	1.22	2.03	1.87	1.59	1.16
19 *	0.91	1.78	2.27	1.78	1.00
20	0.91	1.53	2.27	1.78	1.12
21	1.22	1.53	2.40	2.06	1.16
22 *	1.22	1.79	2.53	2.06	1.08
23	1.53	1.53	2.53	2.16	1.04
24 *	6.73	4.85	4.80	4.32	3.24
25	1.53	2.81	3.33	2.53	1.66
26 *	3.67	4.08	4.93	3.66	2.66
27	2.45	2.81	3.73	2.53	2.35

CBT Item Number	Far Below Basic	Below Basic	Basic	Proficient	Advanced
28 *	3.98	4.08	4.93	3.19	3.20
29	2.45	3.32	4.27	2.72	2.55
30 *	2.45	3.57	4.53	3.00	2.63
31 *	1.68	2.45	0.85	1.26	0.52
32 *	1.68	2.72	0.99	1.26	0.56
33 *	1.68	2.72	0.99	1.26	0.56
34 *	1.68	2.72	0.99	1.26	0.56
35 *	2.02	2.72	1.13	1.46	0.72
36	2.71	2.19	1.84	1.36	0.88
37	2.37	2.19	1.84	1.36	0.84
38	2.72	2.19	2.12	1.46	0.92
39 *	2.73	2.19	2.12	1.75	1.00
40	3.07	2.47	2.26	1.56	1.12
41	3.41	2.19	2.26	1.46	1.08
42	3.07	2.49	2.12	1.56	1.04
43	2.73	2.49	2.26	1.76	1.04
44	3.07	2.21	2.26	1.95	1.12
45	2.73	2.21	2.41	2.05	1.16
46	2.73	2.21	2.41	1.96	1.12
47	2.73	2.21	2.69	2.15	1.20
48 *	3.77	2.21	2.84	2.45	1.72
49 *	3.77	2.49	2.98	2.26	1.36
50	3.77	2.49	2.70	2.26	1.28
51	3.42	2.49	2.70	2.26	1.28
52	3.42	2.77	2.70	2.36	1.28
53	3.42	2.49	2.70	2.26	1.28
54 *	3.77	3.05	2.99	2.56	1.36
55	4.11	2.77	2.70	2.46	1.28
56	3.77	3.05	2.85	2.65	1.32
57	4.45	3.32	2.71	2.56	1.28
58	4.45	3.60	2.71	2.56	1.36
59	4.11	3.32	2.71	2.56	1.28
60	4.45	3.32	2.43	2.46	1.28

* Technology-enhanced item

Table E.9. CBT Percent Omit for Biology by Performance Level on the Spring 2012 CST for TE and MC Items

CBT Item Number	Far Below Basic	Below Basic	Basic	Proficient	Advanced
1	0.00	0.00	0.28	0.07	0.00
2	0.00	0.00	0.48	0.20	0.08
3 *	0.00	0.27	0.55	0.33	0.00
4	0.00	0.00	0.48	0.13	0.00
5	0.00	0.27	0.48	0.33	0.00
6	0.00	0.27	0.48	0.20	0.00
7	0.00	0.27	0.48	0.20	0.00
8 *	0.00	0.27	0.62	0.53	0.15
9	0.00	0.27	0.55	0.33	0.08
10	0.00	0.27	0.55	0.27	0.08
11 *	0.00	0.27	0.48	0.20	0.23

CBT Item Number	Far Below Basic	Below Basic	Basic	Proficient	Advanced
12	0.00	0.27	0.55	0.13	0.08
13	0.00	0.27	0.55	0.33	0.08
14	0.00	0.27	0.55	0.20	0.08
15 *	0.00	0.54	0.48	0.13	0.08
16	0.00	0.27	0.62	0.40	0.15
17	0.00	0.27	0.55	0.27	0.15
18	0.00	0.27	0.69	0.33	0.15
19	0.37	0.54	0.76	0.40	0.23
20 *	0.37	0.54	0.76	0.47	0.46
21	0.74	0.54	0.69	0.33	0.23
22 *	1.84	0.54	1.31	0.60	0.46
23	0.37	0.27	0.83	0.33	0.38
24	0.37	0.27	0.76	0.33	0.46
25	0.74	0.27	0.83	0.40	0.54
26 *	0.74	0.27	0.96	0.53	0.38
27 *	0.74	0.81	0.90	0.33	0.38
28	0.74	0.27	1.03	0.53	0.38
29	0.74	0.27	1.10	0.53	0.46
30	0.74	0.27	0.96	0.53	0.54
31	0.74	0.27	1.03	0.47	0.46
32 *	2.29	1.98	2.11	1.53	0.89
33 *	0.38	0.85	1.38	1.88	1.29
34 *	0.38	0.85	1.38	1.88	1.29
35 *	0.38	1.69	1.53	2.36	1.45
36 *	0.38	1.98	1.82	2.78	1.78
37 *	0.38	3.39	2.11	2.78	1.86
38 *	0.38	2.54	2.40	2.92	2.10
39	0.38	0.29	1.32	0.91	0.98
40	0.76	0.57	1.10	1.20	0.73
41 *	0.76	0.57	1.32	0.98	0.65
42	0.38	0.29	1.32	0.99	0.73
43	0.76	0.29	1.32	0.99	0.73
44 *	0.38	0.58	1.39	1.06	0.82
45	1.15	0.58	1.32	1.13	0.74
46	0.38	0.87	1.39	1.06	0.74
47	0.38	0.87	1.39	1.06	0.82
48 *	2.67	2.90	2.50	1.70	0.90
49	0.38	1.45	1.40	1.34	1.07
50	0.76	1.16	1.47	1.34	1.07
51	0.38	1.16	1.47	1.34	1.07
52	0.38	1.45	1.47	1.34	1.07
53	0.76	1.16	1.47	1.27	1.07
54 *	8.40	4.36	4.57	3.26	1.48
55	0.76	1.45	1.69	1.63	1.23
56	0.76	1.16	1.77	1.34	1.31
57	0.76	1.16	1.85	1.34	1.23
58	0.76	1.16	1.85	1.42	1.48
59	0.76	1.16	1.85	1.42	1.40
60	0.38	1.17	1.77	1.49	1.40

* Technology-enhanced item

Subgroup Summary Statistics

Table E.10. Subgroup Score Summary Statistics: Grade Five Science

	CBT Raw Score					2012 Scale Score				
Group	No. *	Mean	SD	Min	Max	Mean	SD	Min	Max	% Prof. + Adv.
Total	8,426	30.9	9.2	0	56	370.2	65.7	193	600	61.4
Male	4,147	31.4	9.4	0	56	373.8	67.1	193	600	63.7
Female	4,279	30.3	8.9	0	55	366.7	64.0	201	600	59.1
Gender unknown	0	–	–	–	–	–	–	–	–	–
American Indian	52	†32.3	†8.8	†16	†51	†374.5	†67.3	†232	†504	†61.5
Asian American	1,049	36.7	8.9	11	56	411.5	68.9	193	600	83.5
Pacific Islander	41	†30.4	†9.1	†13	†47	†365.7	†56.3	†242	†461	†68.3
Filipino	259	33.5	8.9	1	53	390.8	55.1	226	600	79.2
Hispanic	4,426	27.6	8.1	0	54	347.4	57.4	193	600	47.2
African American	425	28.9	8.2	9	50	356.1	60.3	214	600	53.6
White	2,014	34.9	8.6	8	55	397.7	61.9	201	600	79.5
Two or more races	160	32.7	8.4	9	53	388.5	57.3	252	600	71.9
Ethnicity unknown	0	–	–	–	–	–	–	–	–	–
No special education services	8,028	31.0	9.2	0	56	371.3	65.5	193	600	62.2
Special education services	398	28.3	9.4	10	52	348.8	65.8	220	600	45.5
Special education services unknown	0	–	–	–	–	–	–	–	–	–
English only	4,442	32.6	9.0	0	55	382.8	63.6	201	600	70.3
Initially fluent English proficient	436	35.8	9.1	14	56	411.2	69.0	242	600	81.4
English learner	2,053	24.4	7.3	0	55	321.4	50.0	193	548	26.5
Reclassified fluent English proficient	1,490	33.0	7.9	0	56	387.9	54.0	220	600	77.0
English proficiency unknown	5	–	–	–	–	–	–	–	–	–
Not economically disadvantaged	3,077	35.8	8.6	1	56	406.3	61.7	193	600	83.4
Economically disadvantaged	5,341	28.0	8.3	0	56	349.4	58.4	193	600	48.8
Economic status unknown	8	–	–	–	–	–	–	–	–	–
Graduate school/Postgraduate training	896	37.6	8.9	1	56	422.9	64.9	220	600	87.6
College graduate	1,357	34.8	8.7	6	56	398.9	61.5	193	600	79.5
Some college (includes AA degree)	1,713	31.5	8.2	8	53	372.3	57.4	193	600	66.3
High school graduate	1,783	28.3	8.3	5	53	352.2	57.9	193	600	49.4
Not a high school graduate	1,552	26.4	7.9	6	51	337.9	57.8	201	600	41.4
Parent education level unknown	1,125	29.8	9.4	0	54	363.6	65.2	214	600	57.8
Metropolitan	8,200	30.9	9.2	0	56	370.4	65.8	193	600	61.4
Small town/Rural	226	30.6	8.5	9	49	362.8	62.0	214	548	61.1

*Statistics are based on the subset of CBT records that could be matched to 2012 data and have a valid 2012 score.

† Statistics are based on fewer than 100 records.

Table E.11. Subgroup Score Summary Statistics: Grade Eight Science

	CBT Raw Score					2012 Scale Score				
Group	No. *	Mean	SD	Min	Max	Mean	SD	Min	Max	% Prof. + Adv.
Total	4,863	33.1	10.8	1	60	409.1	97.3	150	600	72.2
Male	2,418	34.1	11.1	1	60	415.2	100.9	150	600	74.2
Female	2,444	32.1	10.4	2	59	403.1	93.2	151	600	70.1
Gender unknown	1	–	–	–	–	–	–	–	–	–
American Indian	45	26.2†	10.6†	8†	51†	360.3†	95.9†	161†	600†	53.3†
Asian American	568	37.8	10.7	2	60	454.7	94.1	179	600	86.4
Pacific Islander	21	†28.4	†11.7	†8	†50	†378.0	†101.7	†195	†559	†57.1
Filipino	134	39.7	9.9	6	59	464.6	93.2	179	600	89.6
Hispanic	2,401	29.9	9.9	4	58	383.8	92.2	150	600	63.1
African American	246	28.8	10.2	5	57	372.1	94.4	161	600	57.7
White	1,348	37.1	10.0	1	59	436.3	89.8	150	600	83.8
Two or more races	100	35.6	11.3	14	59	436.1	106.2	225	600	76.0
Ethnicity unknown	0	–	–	–	–	–	–	–	–	–
No special education services	4,658	33.4	10.7	1	60	411.7	96.2	150	600	73.4
Special education services	205	27.2	10.9	8	56	349.9	102.8	161	600	43.4
Special education services unknown	0	–	–	–	–	–	–	–	–	–
English only	2,616	34.4	10.8	1	60	417.8	96.6	150	600	75.6
Initially fluent English proficient	445	35.7	10.1	6	58	436.9	92.8	195	600	81.3
English learner	630	23.9	8.3	4	56	323.3	81.0	151	600	33.3
Reclassified fluent English proficient	1,170	34.2	9.7	2	59	424.8	83.5	170	600	81.8
English proficiency unknown	2	–	–	–	–	–	–	–	–	–
Not economically disadvantaged	1,872	37.2	10.4	1	60	441.8	91.4	150	600	85.0
Economically disadvantaged	2,988	30.5	10.2	2	59	388.5	95.2	150	600	64.1
Economic status unknown	3	–	–	–	–	–	–	–	–	–
Graduate school/Postgraduate training	536	39.8	10.3	9	60	464.3	88.8	170	600	90.3
College graduate	886	37.6	10.2	2	59	446.6	89.6	170	600	86.3
Some college (includes AA degree)	979	32.6	10.3	1	59	403.8	90.6	150	600	72.0
High school graduate	1,082	30.7	10.2	4	58	391.1	94.5	150	600	65.4
Not a high school graduate	836	29.1	9.5	5	58	374.7	90.6	151	600	60.2
Parent education level unknown	544	31.0	10.8	8	59	391.7	103.1	150	600	63.2
Metropolitan	4,585	33.2	10.8	1	60	410.9	97.5	150	600	72.8
Small town/Rural	278	31.1	10.2	8	58	379.6	88.3	161	600	62.2

* Statistics are based on the subset of CBT records that could be matched to 2012 data and have a valid 2012 score.

† Statistics are based on fewer than 100 records.

Table E.12. Subgroup Score Summary Statistics: Biology

		CBT Raw Score				2012 Scale Score				
Group	No. *	Mean	SD	Min	Max	Mean	SD	Min	Max	% Prof. + Adv.
Total	4,587	29.6	10.8	0	60	360.1	54.3	206	600	57.3
Male	2,270	29.6	11.2	0	59	360.9	57.3	206	600	58.5
Female	2,317	29.5	10.4	2	60	359.4	51.2	228	600	56.2
Gender unknown	0	–	–	–	–	–	–	–	–	–
American Indian	59	†28.0	†10.5	†8	†53	†355.2	†61.5	†240	†529	†50.8
Asian American	333	32.3	10.9	10	59	378.7	60.7	240	600	67.3
Pacific Islander	32	†29.8	†9.6	†15	†53	†360.4	†51.4	†255	†454	†59.4
Filipino	128	32.8	10.2	7	52	374.5	43.7	273	529	70.3
Hispanic	2,358	27.2	10.1	0	58	346.9	49.9	206	600	47.3
African American	266	27.9	10.1	5	55	352.5	48.7	245	529	51.5
White	1,323	33.1	10.9	7	60	378.5	55.0	240	600	71.6
Two or more races	88	†32.9	†9.9	†13	†52	†375.3	†49.9	†240	†492	†76.1
Ethnicity unknown	0	–	–	–	–	–	–	–	–	–
No special education services	4,396	29.9	10.7	2	60	362.0	53.8	206	600	58.8
Special education services	191	21.8	9.3	0	48	316.0	47.1	228	470	22.5
Special education services unknown	0	–	–	–	–	–	–	–	–	–
English only	2,579	30.9	10.9	5	60	367.8	55.1	221	600	63.1
Initially fluent English proficient	307	31.3	10.3	6	53	369.1	53.8	234	600	66.8
English learner	475	19.9	7.6	0	46	308.1	39.7	206	470	13.7
Reclassified fluent English proficient	1,224	30.0	9.6	2	55	361.9	46.0	221	600	59.6
English proficiency unknown	2	–	–	–	–	–	–	–	–	–
Not economically disadvantaged	1,910	32.6	10.7	7	60	377.7	55.4	234	600	70.8
Economically disadvantaged	2,675	27.4	10.2	0	58	347.6	49.9	206	529	47.7
Economic status unknown	2	–	–	–	–	–	–	–	–	–
Graduate school/Postgraduate training	418	36.0	11.1	9	58	399.7	58.4	250	600	83.3
College graduate	741	32.8	10.4	7	59	376.5	50.5	245	600	72.9
Some college (includes AA degree)	1,122	30.2	10.4	5	56	362.2	50.9	221	600	59.7
High school graduate	1,024	27.0	9.6	8	53	346.8	49.2	234	600	46.4
Not a high school graduate	797	25.8	9.8	0	56	339.3	47.4	221	529	41.5
Parent education level unknown	485	29.4	11.3	8	60	358.7	58.7	206	529	54.6
Metropolitan	4,297	29.9	10.8	0	60	361.3	54.3	206	600	58.1
Small town/Rural	290	24.2	9.3	6	53	342.8	52.0	250	600	46.2

* Statistics are based on the subset of CBT records that could be matched to 2012 data and have a valid 2012 score.

† Statistics are based on fewer than 100 records.

DIF Results

Table E.13. Complete DIF Results for Grade Six: Gender, English Proficiency, Special Services, Economic Status, Area, and Parent Education Comparisons

	Male/Female		English Speaker/ English Learner		No Special Services/ Special Services		Not Econ. Disadv./ Econ. Disadv.		Metropolitan/ Small Town-Rural		High Parent Ed. 1/ Low Parent Ed. 1 [a]		High Parent Ed. 2/ Low Parent Ed. 2 [b]	
Item	Cat.	MH D-DIF	Cat.	MH D-DIF	Cat.	MH D-DIF	Cat.	MH D-DIF	Cat.	MH D-DIF	Cat.	MH D-DIF	Cat.	MH D-DIF
1	A-	–0.98	A-	–0.24	A+	0.39	A-	–0.28	A-	–0.32	A-	–0.14	A+	0.00
2	A+	0.15	A-	–0.54	A-	–0.34	A-	–0.76	A+	0.39	A-	–0.32	A-	–0.64
3 *	A-	–0.92	A-	–0.38	A-	-1.06	A-	–0.88	A-	–0.85	A-	–0.71	A-	–0.57
4	A-	–0.05	A-	–0.34	A+	0.36	A-	–0.11	A+	0.61	A-	–0.07	A-	–0.03
5	B-	-1.11	A-	–0.15	A+	0.21	A-	–0.39	B+	1.51	A-	–0.02	A-	–0.33
6	A+	0.21	A-	–0.93	A-	–0.42	A-	–0.82	A-	–0.58	A-	–0.80	A-	–0.75
7	B-	-1.17	A-	–0.14	A+	0.91	A-	–0.71	C+	1.62	A-	–0.18	A-	–0.55
8	A+	0.22	A+	0.00	A+	0.13	A+	0.13	A-	–0.32	A+	0.03	A-	–0.10
9	A+	0.12	A-	–0.08	A-	–0.19	A+	0.32	B+	1.08	A+	0.37	A+	0.27
10*	A+	0.26	A-	–0.54	A-	–0.51	A-	–0.67	A+	0.39	A-	–0.75	A-	–0.50
11	A-	–0.05	A+	0.27	A+	0.09	A-	–0.35	A+	0.31	A-	–0.38	A-	–0.45
12	A-	–0.32	A-	–0.44	A+	0.05	A-	–0.06	A+	0.68	A-	–0.23	A-	–0.04
13	A+	0.45	A-	–0.08	A+	0.00	A-	–0.03	A+	0.11	A-	–0.33	A+	0.00
14 *	A-	–0.17	A+	0.03	A-	–0.51	A-	–0.01	B-	-1.22	A-	–0.41	A-	–0.58
15	A+	0.08	A-	–0.34	A+	0.04	A-	–0.29	A-	–0.26	A-	–0.44	A-	–0.35
16 *	A-	–0.33	A-	–0.08	A+	0.32	A+	0.02	A-	–0.82	A+	0.05	A-	–0.21
17 *	A-	–0.27	A-	–0.21	A+	0.03	A-	–0.40	A+	0.18	A+	0.17	A-	–0.04
18 *	A+	0.55	A-	–0.25	A-	–0.96	A+	0.11	A-	-1.01	A+	0.70	A+	0.00
19	A+	0.17	A-	–0.72	A-	–0.23	A-	–0.56	A-	–0.47	A-	–0.29	A-	–0.32
20	A-	–0.18	A-	–0.82	A+	0.08	A+	0.31	A-	–0.44	A+	0.25	A+	0.18
21	A-	–0.08	A-	–0.78	A+	0.07	A+	0.01	B+	1.07	A+	0.17	A+	0.17
22 *	A+	0.50	A-	–0.09	A+	0.49	A+	0.03	A-	–0.47	A-	–0.37	A-	–0.40
23	A-	–0.18	A-	–0.28	A+	0.11	A-	–0.56	A+	0.14	A-	–0.36	A-	–0.50
24 *	A-	–0.83	A+	0.02	A-	–0.18	A-	–0.15	A+	0.41	A-	–0.26	A-	–0.18
25	A+	0.01	A-	–0.37	A+	0.76	A-	–0.23	A+	0.79	A-	–0.05	A-	–0.17
26	A-	–0.28	A-	–0.47	A+	0.41	A-	–0.17	A+	0.49	A-	–0.40	A-	–0.30
27 *	A-	–0.78	A+	0.14	A+	0.21	A-	–0.34	A+	0.14	A-	–0.20	A-	–0.11
28	A-	–0.31	A+	0.02	A+	0.09	A-	–0.06	A+	0.67	A-	–0.14	A+	0.00
29	A+	0.53	A-	–0.99	A-	–0.12	A-	–0.40	A+	0.28	A-	–0.15	A-	–0.25
30 *	B-	-1.35	A+	0.00	A-	–0.04	A+	0.24	A-	–0.09	A+	0.35	A+	0.24
31 *	A-	–0.98	A-	–0.88	A+	0.30	A+	0.27	A-	-1.66	A+	0.64	A-	–0.51
32 *	A+	0.58	A-	–0.30	A-	–0.44	A+	0.26	A-	–0.76	A+	0.15	A+	0.14
33 *	A-	–0.23	A-	–0.32	A-	–0.22	A+	0.31	A-	–0.44	A+	0.37	A+	0.02
34 *	A+	0.03	A-	–0.29	A+	0.04	A-	–0.04	A-	–0.27	A-	–0.04	A+	0.00
35 *	A-	–0.36	A-	–0.32	A-	–0.14	A-	–0.28	A-	–0.30	A-	–0.29	A-	–0.03
36	A-	–0.25	A-	–0.06	A+	0.03	A-	–0.24	A-	–0.57	A+	0.11	A-	–0.08
37	A-	–0.14	A+	0.01	A-	–0.10	A+	0.43	A-	–0.52	A+	0.29	A-	–0.17
38 *	A-	–0.21	A-	–0.10	A-	–0.06	A-	–0.35	A+	0.05	A-	–0.35	A-	–0.11

Item	Male/Female		English Speaker/ English Learner		No Special Services/ Special Services		Not Econ. Disadv./ Econ. Disadv.		Metropolitan/ Small Town-Rural		High Parent Ed. 1/ Low Parent Ed. 1 [a]		High Parent Ed. 2/ Low Parent Ed. 2 [b]	
	Cat.	MH D-DIF	Cat.	MH D-DIF	Cat.	MH D-DIF	Cat.	MH D-DIF	Cat.	MH D-DIF	Cat.	MH D-DIF	Cat.	MH D-DIF
39	A-	–0.26	A-	–0.18	A+	0.06	A-	–0.32	A-	–0.29	A+	0.06	A-	–0.26
40	A-	–0.90	A-	–0.50	A+	0.22	A-	–0.36	A-	–0.21	A-	–0.18	A-	–0.39
41	A+	0.96	A-	–0.28	A-	–0.95	A-	–0.72	A-	–0.03	A-	–0.54	A-	–0.73
42	A-	–0.39	A-	–0.29	A+	0.13	A-	–0.20	A-	–0.61	A-	–0.22	A-	–0.28
43	A+	0.13	A-	–0.37	A-	–0.35	A-	–0.31	A-	–0.03	A-	–0.04	A-	–0.26
44	A+	0.45	A+	0.05	A+	0.40	A+	0.03	A-	–0.28	A+	0.10	A+	0.23
45 *	A+	0.60	A+	0.14	A-	–0.06	A+	0.22	A-	–0.33	A-	–0.06	A+	0.24
46	A+	0.22	A-	–0.12	A+	0.18	A+	0.26	A+	0.12	A+	0.13	A-	–0.07
47	A+	0.75	A-	–0.40	A-	–0.16	A-	–0.42	A-	–0.23	A-	–0.46	A-	–0.28
48	A+	0.20	A-	–0.62	A+	0.19	A-	–0.54	A-	–0.64	A-	–0.27	A-	–0.57
49 *	A-	–0.56	A-	–0.19	A+	0.42	A-	–0.37	A-	–0.36	A-	–0.06	A-	–0.14
50	A-	–0.64	A-	–0.57	A+	0.22	A-	–0.86	A+	0.21	A-	–0.51	A-	–0.71
51	A-	–0.28	A-	–0.51	A-	–0.02	A-	–0.42	A+	0.85	A-	–0.41	A-	–0.54
52 *	A+	0.32	A-	–0.21	A-	–0.26	B-	-1.01	A-	–0.62	A-	–0.61	A-	–0.54
53	A+	0.78	A-	–0.07	A+	0.36	A-	–0.34	A+	0.23	A-	–0.36	A-	–0.30
54 *	A+	0.33	A+	0.18	A-	–0.70	A+	0.28	A-	–0.23	A+	0.21	A+	0.30
55	A+	0.47	A-	–0.46	A-	–0.23	A-	–0.45	A+	0.19	A-	–0.43	A-	–0.27
56	A-	–0.26	A+	0.28	A+	0.42	A-	–0.27	A-	–0.62	A+	0.16	A+	0.20
57	A+	0.69	A+	0.07	A-	–0.32	A+	0.20	A+	0.43	A+	0.16	A-	–0.04
58	A+	0.60	A-	–0.40	A-	–0.37	A-	–0.32	A+	0.05	A-	–0.08	A-	–0.11
59	A-	–0.27	A+	0.09	A-	–0.27	A-	–0.46	A-	–0.04	A-	–0.42	A-	–0.26
60	A-	–0.66	A-	–0.06	A+	0.94	A-	–0.25	A+	0.12	A-	–0.46	A-	–0.33

* Technology-enhanced item

[a] High Parent Ed. 1 = College Graduate and above; Low Parent Ed. 1 = Some College and below

[b] High Parent Ed. 2 = Some College and above; Low Parent Ed. 2 = High School Graduate and below

Table E.14. Complete DIF Results for Grade Five Science: Ethnicity Comparisons

Item	White/ American Indian †		White/Asian American		White/Pacific Islander †		White/ Filipino		White/ Combined Asian		White/ Hispanic		White/ African American	
	Cat.	MH D-DIF	Cat.	MH D-DIF	Cat.	MH D-DIF	Cat.	MH D-DIF	Cat.	MH D-DIF	Cat.	MH D-DIF	Cat.	MH D-DIF
1	–	–	A+	0.37	–	–	A-	–0.13	A+	0.26	A-	–0.10	A+	0.23
2	–	–	A-	–0.61	–	–	A-	–0.36	A-	–0.50	B-	–1.24	A-	–0.12
3 *	–	–	A+	0.41	–	–	A+	0.14	A+	0.34	B-	–1.26	A-	–0.25
4	–	–	A-	–0.18	–	–	A+	0.04	A-	–0.11	A-	–0.16	A-	–0.11
5	–	–	A-	–0.29	–	–	A-	–0.76	A-	–0.43	A-	–0.57	C-	–2.14
6	–	–	A-	–0.26	–	–	A-	–0.94	A-	–0.45	B-	–1.06	A-	–0.50
7	–	–	A-	–0.71	–	–	A-	–0.58	A-	–0.68	B-	–1.07	A-	–0.67
8	–	–	A+	0.12	–	–	A-	–0.79	A-	–0.13	A-	–0.03	A-	–0.48
9	–	–	A-	–0.40	–	–	A-	–0.11	A-	–0.30	A+	0.08	A-	–0.75
10 *	–	–	A+	0.37	–	–	A+	0.14	A+	0.28	B-	–1.01	B-	–2.12
11	–	–	A+	0.72	–	–	A+	0.59	A+	0.66	A-	–0.37	A-	–0.51

Item	White/ American Indian †		White/Asian American		White/Pacific Islander †		White/ Filipino		White/ Combined Asian		White/ Hispanic		White/ African American	
	Cat.	MH D-DIF	Cat.	MH D-DIF	Cat.	MH D-DIF	Cat.	MH D-DIF	Cat.	MH D-DIF	Cat.	MH D-DIF	Cat.	MH D-DIF
12	–	–	A+	0.15	–	–	A+	0.22	A+	0.19	A+	0.01	A-	–0.12
13	–	–	A+	0.50	–	–	A-	–0.04	A+	0.36	A+	0.08	A+	0.18
14*	–	–	A+	0.32	–	–	A+	0.44	A+	0.32	A-	–0.31	A-	–0.06
15	–	–	A+	0.44	–	–	A-	–0.29	A+	0.25	A-	–0.47	A+	0.08
16 *	–	–	A+	0.11	–	–	A+	0.09	A+	0.07	A-	–0.01	A-	–0.37
17 *	–	–	A+	0.27	–	–	A-	–1.19	A-	–0.28	A+	0.24	A-	–0.31
18 *	–	–	A+	0.88	–	–	B-	–1.54	A+	0.14	A+	0.04	A+	0.84
19	–	–	A-	–0.40	–	–	A-	–0.35	A-	–0.31	A-	–0.70	A-	–0.41
20	–	–	A+	0.86	–	–	A+	0.26	A+	0.71	A+	0.21	A-	–0.04
21	–	–	A+	0.47	–	–	A-	–0.45	A+	0.18	A-	–0.40	A+	0.07
22 *	–	–	A+	0.96	–	–	A+	0.47	A+	0.86	A-	–0.02	B-	–1.35
23	–	–	A-	–0.63	–	–	A-	–0.61	A-	–0.63	A-	–0.90	A-	–0.17
24 *	–	–	A+	0.28	–	–	A+	0.18	A+	0.18	A-	–0.48	A-	–0.24
25	–	–	B+	1.03	–	–	A-	–0.03	A+	0.72	A-	–0.18	A+	0.56
26	–	–	A+	0.04	–	–	A-	–0.34	A-	–0.03	A-	–0.24	A+	0.23
27 *	–	–	A+	0.21	–	–	A-	–0.52	A+	0.03	A-	–0.55	A-	–0.89
28	–	–	A+	0.49	–	–	A-	–0.27	A+	0.26	A-	–0.06	A-	–0.07
29	–	–	A-	–0.87	–	–	B-	–1.05	A-	–0.93	A-	–0.75	A-	–0.88
30 *	–	–	A+	0.61	–	–	A-	–0.36	A+	0.44	A+	0.13	A-	–0.15
31 *	–	–	A-	–1.05	–	–	A+	1.73	A-	–0.25	A-	–0.03	A-	–1.19
32 *	–	–	A+	0.45	–	–	B+	1.19	A+	0.68	A+	0.36	A+	0.53
33 *	–	–	A-	–0.02	–	–	A-	–0.32	A-	–0.13	A-	–0.11	A-	–0.70
34 *	–	–	A-	–0.12	–	–	A-	–0.13	A-	–0.05	A-	–0.16	A+	0.14
35 *	–	–	A+	0.13	–	–	A-	–0.51	A-	–0.02	A-	–0.39	A-	–0.21
36	–	–	A-	–0.21	–	–	A-	–0.50	A-	–0.25	A-	–0.14	A-	–0.43
37	–	–	A-	–0.56	–	–	B-	–1.15	A-	–0.75	A-	–0.04	A-	–0.16
38 *	–	–	A+	0.05	–	–	A+	0.04	A+	0.06	A-	–0.32	A-	–0.19
39	–	–	A-	–0.32	–	–	A-	–0.01	A-	–0.25	A-	–0.40	A-	–0.62
40	–	–	A-	–0.68	–	–	A-	–0.39	A-	–0.65	A-	–0.57	A-	–0.22
41	–	–	A+	0.29	–	–	A+	0.26	A+	0.32	A-	–0.46	A-	–0.23
42	–	–	A+	0.15	–	–	A-	–0.18	A+	0.09	A-	–0.07	A-	–0.26
43	–	–	A+	0.23	–	–	A-	–0.56	A-	–0.03	A-	–0.34	A-	–0.56
44	–	–	A+	0.92	–	–	A+	0.70	A+	0.82	A+	0.36	A+	0.03
45 *	–	–	A-	–0.07	–	–	A+	0.13	A-	–0.04	A+	0.16	A-	–0.27
46	–	–	A+	0.05	–	–	A+	0.17	A+	0.05	A+	0.11	A+	0.22
47	–	–	A+	0.84	–	–	A+	0.85	A+	0.80	A-	–0.30	A-	–0.33
48	–	–	A+	0.19	–	–	A+	0.95	A+	0.36	A-	–0.58	A-	–0.54
49 *	–	–	A+	0.25	–	–	A-	–0.15	A+	0.14	A-	–0.33	A-	–0.13
50	–	–	A-	–0.14	–	–	A+	0.84	A+	0.04	A-	–0.58	B-	–1.07
51	–	–	A+	0.11	–	–	A+	0.20	A+	0.16	A-	–0.84	A-	–0.90
52 *	–	–	A+	0.61	–	–	A+	0.66	A+	0.57	A-	–0.64	A+	0.03
53	–	–	A-	–0.48	–	–	B-	–1.05	A-	–0.54	A-	–0.30	A-	–0.92
54 *	–	–	A+	0.08	–	–	A-	–0.17	A-	–0.01	A+	0.38	A+	0.02

Item	White/ American Indian †		White/Asian American		White/Pacific Islander †		White/ Filipino		White/ Combined Asian		White/ Hispanic		White/ African American	
	Cat.	MH D-DIF	Cat.	MH D-DIF	Cat.	MH D-DIF	Cat.	MH D-DIF	Cat.	MH D-DIF	Cat.	MH D-DIF	Cat.	MH D-DIF
55	–	–	A-	–0.02	–	–	A-	–0.49	A-	–0.16	A-	–0.46	A-	–0.48
56	–	–	A+	0.17	–	–	A-	–0.44	A+	0.11	A-	–0.29	A-	–0.84
57	–	–	A+	0.12	–	–	B+	1.29	A+	0.41	A+	0.31	A+	0.17
58	–	–	A+	0.05	–	–	A+	0.28	A+	0.13	A+	0.00	A+	0.24
59	–	–	A+	0.31	–	–	A-	–0.23	A+	0.15	A-	–0.50	A-	–0.36
60	–	–	A+	0.28	–	–	A-	–0.72	A-	–0.02	A-	–0.43	A-	–0.72

* Technology-enhanced item

† Insufficient sample size for DIF analysis

Table E.15. Complete DIF Results for Grade Eight Science: Gender, English Proficiency, Economic Status, Area, and Parent Education Comparisons

Item	Male/Female		English Speaker/ English Learner		No Special Services/ Special Services		Not Econ. Disadv./ Econ. Disadv.		Metropolitan/ Small Town-Rural		High Parent Ed. 1/Low Parent Ed. 1 [a]		High Parent Ed. 2/Low Parent Ed 2 [b]	
	Cat.	MH D-DIF	Cat.	MH D-DIF	Cat.	MH D-DIF	Cat.	MH D-DIF	Cat.	MH D-DIF	Cat.	MH D-DIF	Cat.	MH D-DIF
1	A+	0.67	A-	–0.11	A-	–0.58	A-	–0.43	A-	–0.37	A-	–0.52	A-	–0.19
2	A-	–0.60	A+	0.04	A-	–0.77	A-	–0.09	A-	–0.40	A-	–0.13	A-	–0.18
3 *	A-	–0.06	A-	–0.76	A-	–0.96	A-	–0.66	A-	–0.45	A-	–0.61	A-	–0.41
4	A+	0.16	A+	0.33	A-	–0.66	A+	0.18	A-	–0.70	A+	0.06	A-	–0.03
5	A-	–0.27	A-	–0.63	A+	0.03	A-	–0.70	A+	0.65	A-	–0.55	A-	–0.71
6	A+	0.15	A+	0.04	A-	–0.70	A-	–0.18	A+	0.27	A-	–0.48	A-	–0.11
7	B+	1.17	A-	–0.77	B-	–1.25	A+	0.03	A+	0.45	A-	–0.23	A+	0.25
8 *	A-	–0.65	A-	–0.65	A-	–0.65	A-	–0.72	B+	1.33	A-	–0.54	A-	–0.68
9	A-	–0.18	A-	–0.50	A+	0.05	A-	–0.09	A+	0.19	A-	–0.23	A-	–0.20
10	A-	–0.22	A+	0.06	A+	0.12	A-	–0.40	A+	0.12	A-	–0.43	A-	–0.23
11 *	A+	0.23	A-	–0.72	A+	0.30	A-	–0.73	A-	–0.12	A-	–0.46	A-	–0.55
12	A+	0.53	A+	0.08	A+	0.08	A-	–0.60	A+	0.28	A-	–0.80	A-	–0.58
13	A+	0.72	A+	0.16	A-	–0.35	A-	–0.40	A-	–0.39	A-	–0.05	A-	–0.12
14	A+	0.48	A+	0.42	A+	0.52	A-	–0.09	A-	–0.39	A-	–0.13	A+	0.14
15	A-	–0.61	A-	–0.24	A+	0.27	A-	–0.60	A+	0.49	A-	–0.68	A-	–0.66
16 *	A-	–0.59	A-	–0.32	A+	0.26	A-	–0.40	A+	0.47	A-	–0.38	A-	–0.41
17	A-	–0.30	A+	0.10	A-	–0.56	A-	–0.58	A+	0.54	A-	–0.72	A-	–0.72
18	A+	0.31	A-	–0.34	A-	–0.17	A-	–0.16	A+	0.11	A-	–0.09	A-	–0.09
19 *	A-	–0.55	A-	–0.57	A-	–0.78	A-	–0.41	C+	1.76	A-	–0.74	A-	–0.59
20	A+	0.17	A-	–0.22	A-	–0.56	A+	0.54	A+	0.08	A+	0.13	A+	0.19
21	A+	0.02	B-	–1.27	A-	–0.12	A-	–0.12	B+	1.10	A-	–0.11	A-	–0.36
22 *	A-	–0.96	A+	0.02	A+	0.46	A-	–0.25	A+	0.85	A-	–0.20	A-	–0.12
23	A-	–0.54	A-	–0.23	A-	–0.31	A-	–0.29	A+	0.14	A-	–0.80	A-	–0.32
24 *	A-	–0.44	A-	–0.26	B-	–1.54	A+	0.06	B+	1.11	A+	0.10	A+	0.22
25	A+	0.13	A-	–0.26	A-	–0.23	A-	–0.20	A-	–0.33	A-	–0.43	A-	–0.15
26 *	A+	0.17	A-	–0.31	A-	–1.00	A-	–0.70	A+	0.73	A-	–0.72	A-	–0.42
27	A+	0.09	A-	–0.81	A+	0.03	B-	–1.28	A-	–0.08	B-	–1.20	B-	–1.03

	Male/Female		English Speaker/ English Learner		No Special Services/ Special Services		Not Econ. Disadv./ Econ. Disadv.		Metropolitan/ Small Town-Rural		High Parent Ed. 1/Low Parent Ed. 1 [a]		High Parent Ed. 2/Low Parent Ed 2 [b]	
Item	Cat.	MH D-DIF	Cat.	MH D-DIF	Cat.	MH D-DIF	Cat.	MH D-DIF	Cat.	MH D-DIF	Cat.	MH D-DIF	Cat.	MH D-DIF
28 *	A-	–0.68	A+	0.20	A+	0.07	A-	–0.80	A+	0.69	B-	–1.15	A-	–1.00
29	A-	–0.09	A-	–0.77	A-	–0.55	A-	–0.38	A+	0.39	A-	–0.41	A-	–0.44
30 *	A+	0.92	A-	–0.15	A-	–0.70	A-	–0.88	A+	0.01	A-	–0.63	A-	–0.55
31 *	A+	0.37	C-	–2.00	A-	–0.96	A-	–0.68	A+	0.23	A-	–0.45	A-	–0.51
32 *	A-	–0.87	A-	–0.15	A+	0.03	A-	–0.76	A+	1.00	A-	–0.36	A-	–0.69
33 *	A+	0.28	A-	–0.99	A-	–0.01	A-	–0.65	A+	0.58	A-	–0.29	A-	–0.27
34 *	A-	–0.86	B+	1.99	A-	–0.15	A+	0.23	A+	0.03	A-	–0.47	A+	0.56
35 *	A-	–0.28	B-	–1.10	B-	–1.98	A-	–0.51	A-	–0.15	A-	–0.73	A-	–0.56
36	A+	0.11	A-	–0.30	B-	–1.09	A+	0.13	A+	0.76	A-	–0.29	A-	–0.02
37	A-	–0.33	A-	–0.58	A-	–0.71	A+	0.21	B+	1.27	A+	0.74	A+	0.71
38	A+	0.05	A-	–0.72	A+	0.35	A-	–0.53	A-	–0.24	A-	–0.51	A-	–0.43
39 *	A+	0.01	A-	–0.49	A+	0.51	A-	–0.33	A+	0.00	A-	–0.31	A-	–0.20
40	A-	–0.22	A+	0.28	A+	0.04	A-	–0.30	A+	0.16	A-	–0.55	A-	–0.54
41	A-	–0.03	A-	–0.03	A-	–0.05	A-	–0.47	A+	0.38	A-	–0.60	A-	–0.53
42	B-	–1.00	C-	–1.70	A-	–0.28	A-	–0.87	A+	0.37	A-	–0.97	A-	–0.16
43	A-	–0.72	A-	–0.03	A-	–0.65	A+	0.03	A-	–0.03	A+	0.38	A+	0.32
44	A-	–0.97	A-	–0.32	A+	0.09	A-	–0.32	A+	0.19	A-	–0.37	A-	–0.13
45	A-	–0.33	A-	–0.11	A+	0.14	A-	–0.33	A+	0.22	A-	–0.44	A-	–0.23
46	A+	0.08	B-	–1.09	A-	–0.96	A-	–0.12	A+	0.71	A+	0.38	A+	0.15
47	A+	0.04	A-	–0.13	A+	0.08	A-	–0.32	A+	0.03	A-	–0.29	A-	–0.34
48 *	A-	–0.02	A-	–0.53	A+	0.37	A-	–0.57	A-	–0.65	A-	–0.62	A-	–0.31
49 *	A-	–0.78	B-	–1.10	A+	0.47	A-	–0.17	A+	0.44	A-	–0.39	A-	–0.34
50	B-	–1.30	A-	–0.70	A-	–0.38	A+	0.37	A-	–0.94	A-	–0.02	A+	0.53
51	B-	–1.16	A-	–0.51	A+	0.83	A-	–0.81	A+	0.03	A-	–0.81	A-	–0.87
52	B-	–1.05	A-	–0.82	A-	–0.24	A-	–0.05	A+	0.29	A+	0.11	A-	–0.11
53	B-	–1.50	A-	–0.45	A+	0.26	A-	–0.19	A+	0.43	A-	–0.27	A-	–0.23
54 *	A+	0.53	B-	–1.45	A+	0.03	A-	–0.68	A-	–0.67	A-	–0.55	A-	–0.25
55	A-	–0.44	A-	–0.69	A+	0.32	A-	–0.49	A+	0.28	A-	–0.22	A-	–0.17
56	A-	–0.04	A-	–1.00	A+	0.37	A-	–0.89	A-	–0.02	A-	–0.78	A-	–0.77
57	A+	0.01	A-	–0.65	A+	0.19	A-	–0.81	A-	–0.02	A-	–0.88	A-	–0.52
58	B+	1.04	A+	0.51	A+	0.10	A+	0.02	A+	0.14	A-	–0.09	A+	0.06
59	A-	–0.39	A-	–0.84	A+	0.18	A-	–0.41	A+	0.22	A-	–0.69	A-	–0.52
60	A+	0.26	A-	–0.74	A-	–0.25	A-	–0.62	A+	0.07	A-	–0.42	A-	–0.40

* Technology-enhanced item

[a] High Parent Ed. 1 = College Graduate and above; Low Parent Ed. 1 = Some College and below

[b] High Parent Ed. 2 = Some College and above; Low Parent Ed. 2 = High School Graduate and below

Table E.16. Complete DIF Results for Grade Eight Science: Ethnicity Comparisons

	White/ American Indian †		White/Asian American		White/Pacific Islander †		White/ Filipino		White/ Combined Asian		White/ Hispanic		White/ African American	
Item	Cat.	MH D-DIF	Cat.	MH D-DIF	Cat.	MH D-DIF	Cat.	MH D-DIF	Cat.	MH D-DIF	Cat.	MH D-DIF	Cat.	MH D-DIF
1	–	–	A+	0.28	–	–	A+	0.03	A+	0.14	A-	–0.56	A-	–0.67
2	–	–	A+	0.11	–	–	A-	–0.33	A-	–0.06	A-	–0.49	A-	–0.26
3 *	–	–	A+	0.05	–	–	A-	–0.40	A-	–0.14	A-	–0.78	A-	–0.76
4	–	–	A-	–0.02	–	–	A+	0.80	A+	0.16	A-	–0.84	A-	–0.39
5	–	–	A-	–0.76	–	–	A-	–0.99	A-	–0.87	A-	–0.84	A-	–0.60
6	–	–	A+	0.78	–	–	A+	0.89	A+	0.48	A-	–0.34	A-	–0.41
7	–	–	A-	–0.91	–	–	A-	–0.28	A-	–0.74	A-	–0.24	B-	–1.59
8*	–	–	B-	–1.04	–	–	A-	–0.69	B-	–1.00	B-	–1.17	B-	–1.67
9	–	–	B-	–1.12	–	–	A-	–0.97	B-	–1.02	A-	–0.31	A-	–0.44
10	–	–	A+	0.10	–	–	A+	0.12	A+	0.07	A-	–0.22	A-	–0.51
11 *	–	–	A-	–0.20	–	–	A+	0.32	A-	–0.14	A-	–0.76	A+	0.22
12	–	–	A-	–0.10	–	–	A-	–0.29	A-	–0.15	A-	–0.24	A-	–0.65
13	–	–	A+	0.33	–	–	A+	0.78	A+	0.44	A-	–0.34	A-	–0.36
14	–	–	A+	0.33	–	–	A+	0.48	A+	0.28	A+	0.05	A-	–0.54
15	–	–	A+	0.00	–	–	A+	0.49	A+	0.17	A-	–0.69	A-	–0.79
16 *	–	–	A+	0.39	–	–	A-	–0.39	A+	0.13	A-	–0.66	A-	–0.55
17	–	–	A-	–0.13	–	–	A-	–0.50	A-	–0.28	A-	–0.81	A-	–0.76
18	–	–	A+	0.39	–	–	A+	0.28	A+	0.31	A-	–0.04	A+	0.38
19 *	–	–	A-	–0.75	–	–	B-	–1.25	A-	–0.93	A-	–0.58	A-	–0.77
20	–	–	A-	–0.19	–	–	A+	0.08	A-	–0.17	A+	0.39	A-	–0.88
21	–	–	B-	–1.02	–	–	A+	0.75	A-	–0.79	A-	–0.01	A-	–0.91
22 *	–	–	A+	0.19	–	–	B+	1.09	A+	0.26	A+	0.11	A-	–0.67
23	–	–	A+	0.10	–	–	A-	–0.51	A-	–0.06	A-	–0.56	B-	–1.21
24 *	–	–	A+	0.87	–	–	A+	0.99	A+	0.87	A+	0.03	A+	0.26
25	–	–	A-	–0.72	–	–	A-	–0.70	A-	–0.61	A-	–0.06	A-	–0.74
26 *	–	–	A-	–0.48	–	–	A-	–0.22	A-	–0.47	A-	–0.88	A-	–0.59
27	–	–	A-	–0.71	–	–	A-	–0.57	A-	–0.70	B-	–1.43	B-	–1.34
28 *	–	–	B+	1.14	–	–	A+	0.51	A+	0.94	A-	–0.98	A-	–0.29
29	–	–	A-	–0.08	–	–	B+	1.87	A+	0.17	A-	–0.58	A-	–0.79
30 *	–	–	A+	0.13	–	–	A+	0.61	A+	0.16	A-	–0.96	B-	–1.06
31 *	–	–	A-	–0.76	–	–	A-	–0.29	A-	–0.74	A-	–0.89	B-	–1.32
32 *	–	–	A+	0.60	–	–	A+	0.46	A+	0.51	B-	–1.07	A-	–0.86
33 *	–	–	A-	–0.97	–	–	A+	0.04	A-	–0.82	A-	–0.70	A-	–0.86
34 *	–	–	A-	–0.65	–	–	A+	0.93	A-	–0.17	B-	–1.18	A-	–1.66
35 *	–	–	B-	–1.08	–	–	A+	0.11	A-	–0.78	C-	–1.57	A-	–0.91
36	–	–	A+	0.46	–	–	A+	0.43	A+	0.47	A-	–0.21	A-	–0.95
37	–	–	B-	–1.26	–	–	A-	–0.01	B-	–1.02	A-	–0.35	B-	–1.58
38	–	–	A-	–0.09	–	–	A+	0.57	A+	0.09	A-	–0.29	A-	–0.63
39 *	–	–	A+	0.24	–	–	A+	0.22	A+	0.17	A-	–0.21	B-	–1.06
40	–	–	A+	0.85	–	–	A+	0.17	A+	0.71	A-	–0.39	A-	–0.33
41	–	–	A+	0.82	–	–	A+	0.46	A+	0.63	A-	–0.32	A-	–0.15
42	–	–	A-	–0.57	–	–	A+	0.18	A-	–0.47	B-	–1.20	A-	–0.30

Item	White/ American Indian †		White/Asian American		White/Pacific Islander †		White/ Filipino		White/ Combined Asian		White/ Hispanic		White/ African American	
	Cat.	MH D-DIF	Cat.	MH D-DIF	Cat.	MH D-DIF	Cat.	MH D-DIF	Cat.	MH D-DIF	Cat.	MH D-DIF	Cat.	MH D-DIF
43	–	–	A+	0.86	–	–	A+	0.30	A+	0.69	A+	0.01	A-	–0.37
44	–	–	A-	–0.36	–	–	A+	0.18	A-	–0.22	A-	–0.78	A-	–0.85
45	–	–	A+	0.37	–	–	A+	0.10	A+	0.29	A-	–0.66	A+	0.06
46	–	–	A+	0.92	–	–	A+	0.66	A+	0.81	A+	0.00	A-	–0.32
47	–	–	A+	0.15	–	–	A-	–0.27	A+	0.02	A-	–0.59	A-	–0.67
48 *	–	–	A+	0.15	–	–	A+	0.10	A+	0.17	A-	–0.47	A-	–0.29
49 *	–	–	A+	0.06	–	–	A+	0.29	A+	0.10	A-	–0.31	A-	–0.48
50	–	–	B-	–1.59	–	–	B-	–1.95	B-	–1.57	A-	–0.21	A+	0.05
51	–	–	C-	–1.58	–	–	B-	–1.35	C-	–1.56	B-	–1.01	A-	–0.70
52	–	–	A-	–0.33	–	–	A-	–0.07	A-	–0.33	A-	–0.07	A-	–0.41
53	–	–	A+	0.51	–	–	A+	1.50	A+	0.64	A-	–0.21	A-	–0.45
54 *	–	–	A-	–0.56	–	–	A+	0.33	A-	–0.47	A-	–0.72	A-	–0.82
55	–	–	A-	–0.78	–	–	B-	–1.07	A-	–0.83	A-	–0.69	A-	–0.62
56	–	–	A-	–0.74	–	–	A+	0.29	A-	–0.46	B-	–1.07	A-	–0.52
57	–	–	A+	0.07	–	–	A+	0.36	A+	0.11	A-	–0.43	A-	–0.85
58	–	–	A+	0.84	–	–	A+	0.81	A+	0.81	A-	–0.27	A+	0.10
59	–	–	A-	–0.33	–	–	A-	–0.44	A-	–0.38	A-	–0.73	A-	–0.44
60	–	–	A-	–0.63	–	–	A-	–0.48	A-	–0.48	A-	–0.66	A-	–0.41

* Technology-enhanced item

† Insufficient sample size for DIF analysis

Table E.17. Complete DIF Results for Biology: Gender, English Proficiency, Economic Status, Area, and Parent Education Comparisons

Item	Male/Female		English Speaker/ English Learner		No Special Services/ Special Services		Not Econ. Disadv./ Econ. Disadv.		Metropolitan/ Small Town-Rural		High Parent Ed. 1/Low Parent Ed. 1 [a]		High Parent Ed. 2/Low Parent Ed. 2 [b]	
	Cat.	MH D-DIF	Cat.	MH D-DIF	Cat.	MH D-DIF	Cat.	MH D-DIF	Cat.	MH D-DIF	Cat.	MH D-DIF	Cat.	MH D-DIF
1	A+	0.63	A-	–0.59	B-	–1.44	A+	0.20	A+	0.26	A-	–0.50	A-	–0.18
2	A+	0.04	A-	–0.49	A-	–0.41	A+	0.02	A-	–0.21	A-	–0.29	A-	–0.53
3 *	A+	0.52	A+	0.45	A-	–0.23	A+	0.30	A+	0.04	A-	–0.08	A+	0.09
4	A+	0.46	A-	–0.11	A-	–0.48	A-	–0.05	A+	0.78	A-	–0.09	A-	–0.01
5	A+	0.18	A-	–0.40	A-	–0.29	A-	–0.08	A+	0.04	A-	–0.17	A+	0.22
6	A-	–0.24	A+	0.03	A-	–0.17	A+	0.11	A+	0.82	A-	–0.11	A+	0.22
7	A-	–0.35	A-	–0.25	A+	0.00	A-	–0.04	A-	–0.41	A+	0.42	A+	0.18
8 *	A+	0.07	A+	0.13	B+	1.25	A-	–0.03	A+	0.38	A-	–0.18	A+	0.02
9	A+	0.52	A+	0.22	C-	–1.85	A-	–0.13	A-	–0.40	A-	–0.29	A-	–0.31
10	A-	–0.34	A-	–0.01	A+	0.05	A-	–0.05	A+	0.63	A-	–0.15	A-	–0.08
11 *	A+	0.29	A+	0.06	A-	–0.08	A+	0.40	C-	–1.78	A+	0.40	A+	0.41
12	A-	–0.26	B-	–1.03	A-	–0.12	A-	–0.47	A+	0.04	A-	–0.32	A-	–0.47
13	A-	–0.01	A+	0.11	A-	–0.25	A-	–0.14	A+	0.03	A-	–0.33	A-	–0.31
14	B+	1.40	A+	0.38	A+	0.42	A+	0.11	A-	–0.35	A+	0.05	A+	0.10
15 *	A-	–0.41	A-	–0.45	A-	–0.22	A+	0.07	A+	0.14	A+	0.34	A+	0.25

Item	Male/Female		English Speaker/ English Learner		No Special Services/ Special Services		Not Econ. Disadv./ Econ. Disadv.		Metropolitan/ Small Town-Rural		High Parent Ed. 1/Low Parent Ed. 1 [a]		High Parent Ed. 2/Low Parent Ed. 2 [b]	
	Cat.	MH D-DIF	Cat.	MH D-DIF	Cat.	MH D-DIF	Cat.	MH D-DIF	Cat.	MH D-DIF	Cat.	MH D-DIF	Cat.	MH D-DIF
16	A+	0.81	A+	0.44	A+	0.56	A-	–0.08	A-	–0.26	A-	–0.23	A+	0.04
17	A+	0.88	A+	0.07	A-	–0.67	A+	0.66	C-	–1.69	A+	0.45	A+	0.26
18	A+	0.31	A+	0.08	A-	–0.25	A+	0.04	A-	–0.70	A+	0.05	A-	–0.09
19	A+	0.89	A-	–0.57	A-	–0.36	A-	–0.33	A-	–0.13	A-	–0.28	A-	–0.33
20 *	A-	–0.56	A+	0.04	A-	–0.41	A-	–0.33	A-	–0.04	A-	–0.53	A-	–0.32
21	A-	–0.06	A-	–0.55	A-	–0.18	A+	0.06	A-	–0.62	A-	–0.10	A+	0.09
22 *	A-	–0.35	A-	–0.50	A-	–0.28	A+	0.07	A-	–0.38	A-	–0.21	A-	–0.02
23	A-	–0.47	A-	–0.41	A+	0.21	A-	–0.13	A-	–0.14	A+	0.05	A-	–0.10
24	A-	–0.44	A-	–0.67	A-	–0.24	A-	–0.44	A-	–0.66	A-	–0.16	A-	–0.12
25	A+	0.26	A-	–0.70	A-	–0.83	A-	–0.29	A-	–0.23	A-	–0.20	A+	0.03
26 *	A-	–0.42	A-	–0.24	A+	0.64	A-	–0.23	A-	–0.68	A-	–0.28	A-	–0.42
27 *	A-	–0.56	A-	–0.88	A+	0.10	A-	–0.69	A-	–0.24	A-	–0.56	A-	–0.62
28	A-	–0.03	A-	–0.38	A-	–0.94	A+	0.19	A-	–0.11	A-	–0.01	A+	0.00
29	A+	0.02	A-	–0.39	A-	–0.46	A-	–0.08	A-	–0.50	A+	0.00	A-	–0.21
30	A+	0.41	A+	0.01	A-	–0.05	A+	0.68	B-	–1.47	A+	0.59	A+	0.44
31	A-	–0.18	A+	0.35	A+	0.16	A+	0.16	A-	–0.98	A+	0.15	A+	0.22
32 *	A+	0.46	B-	–1.49	A-	–0.71	A-	–0.08	A-	–0.76	A+	0.01	A-	–0.59
33 *	A+	0.15	A+	0.03	A-	–0.67	A-	–0.01	C-	–2.49	A+	0.12	A+	0.09
34 *	A-	–0.01	A-	–0.27	A-	–0.21	A-	–0.13	B-	–1.09	A-	–0.15	A-	–0.20
35 *	A-	–0.30	A-	–0.80	A-	–0.57	A-	–0.30	C-	–1.65	A+	0.06	A-	–0.29
36 *	A-	–0.30	A+	0.40	A-	–0.97	A+	0.30	C-	–2.49	A+	0.33	A+	0.05
37 *	A+	0.10	B-	–1.17	A-	–0.47	A-	–0.61	C-	–1.93	A-	–0.38	A-	–0.61
38 *	A+	0.05	A-	–0.27	A-	–0.02	A-	–0.05	A-	–0.72	A-	–0.43	A-	–0.39
39	A+	0.50	A+	0.41	A-	–0.70	A+	0.25	A-	–0.39	A-	–0.15	A+	0.01
40	A+	0.26	A-	–0.90	A-	–0.95	A-	–0.22	A-	–0.95	A-	–0.19	A-	–0.14
41 *	A-	–0.76	A-	–0.42	A+	0.07	A-	–0.19	A+	0.04	A-	–0.30	A-	–0.35
42	A+	0.13	A-	–0.61	A-	–0.34	A-	–0.26	A-	–0.48	A-	–0.03	A+	0.00
43	A-	–0.10	A-	–0.64	A-	–0.64	A-	–0.14	A-	–0.50	A-	–0.14	A-	–0.14
44 *	A+	0.85	A+	0.44	A-	–0.38	A+	0.05	A-	–0.63	A-	–0.10	A-	–0.02
45	A-	–0.07	A-	–0.86	A+	0.10	A-	–0.04	A-	–0.37	A-	–0.20	A-	–0.09
46	A-	–0.38	A-	–0.92	A-	–0.11	A-	–0.30	A-	–0.37	A-	–0.29	A-	–0.40
47	A+	0.35	A-	–0.04	A-	–0.70	A+	0.34	A-	–0.65	A+	0.06	A+	0.11
48 *	A-	–0.97	A-	–0.74	B-	–1.07	A+	0.09	B-	–1.19	A-	–0.26	A-	–0.26
49	A-	–0.60	A-	–0.18	B+	1.05	A-	–0.33	A-	–0.46	A-	–0.49	A-	–0.49
50	A+	0.47	A-	–0.20	A+	0.41	A-	–0.07	B-	–1.18	A+	0.04	A-	–0.12
51	A-	–0.08	A-	–0.67	A+	0.19	A+	0.04	A-	–0.69	A+	0.00	A-	–0.08
52	A+	0.18	A-	–0.75	A+	0.06	A-	–0.48	B-	–1.19	A-	–0.70	A-	–0.70
53	A+	0.31	A-	–0.64	A+	0.15	A-	–0.05	C-	–1.85	A-	–0.18	A-	–0.42
54 *	A-	–0.25	A-	–0.37	A-	–0.29	A-	–0.05	A-	–0.36	A-	–0.03	A-	–0.29
55	A+	0.35	A-	–0.27	A-	–0.18	A-	–0.18	A-	–0.53	A-	–0.16	A-	–0.24
56	A-	–0.30	A-	–0.07	A+	0.62	A+	0.40	A-	–0.11	A+	0.16	A+	0.13

	Male/Female		English Speaker/ English Learner		No Special Services/ Special Services		Not Econ. Disadv./ Econ. Disadv.		Metropolitan/ Small Town-Rural		High Parent Ed. 1/Low Parent Ed. 1 [a]		High Parent Ed. 2/Low Parent Ed. 2 [b]	
Item	Cat.	MH D-DIF	Cat.	MH D-DIF	Cat.	MH D-DIF	Cat.	MH D-DIF	Cat.	MH D-DIF	Cat.	MH D-DIF	Cat.	MH D-DIF
57	A-	–0.38	A-	–0.54	A-	–0.43	A+	0.05	A-	–0.86	A-	–0.25	A-	–0.29
58	A+	0.51	A-	–0.25	A-	–0.47	A+	0.08	A-	–0.92	A-	–0.14	A-	–0.25
59	A+	0.18	A-	–0.27	A+	0.59	A-	–0.42	A-	–0.64	A-	–0.02	A-	–0.10
60	A+	0.24	A-	–0.22	A-	–0.35	A-	–0.07	B-	–1.01	A-	–0.26	A-	–0.07

* Technology-enhanced item

[a] High Parent Ed. 1 = College Graduate and above; Low Parent Ed. 1 = Some College and below

[b] High Parent Ed. 2 = Some College and above; Low Parent Ed. 2 = High School Graduate and below

Table E.18. Complete DIF results for Biology: Ethnicity Comparisons

	White/ American Indian †		White/Asian American		White/Pacific Islander†		White/ Filipino		White/ Combined Asian		White/ Hispanic		White/ African American	
Item	Cat.	MH D-DIF	Cat.	MH D-DIF	Cat.	MH D-DIF	Cat.	MH D-DIF	Cat.	MH D-DIF	Cat.	MH D-DIF	Cat.	MH D-DIF
1	–	–	A+	0.35	–	–	A+	3.19	A+	0.66	A+	0.09	A+	0.09
2	–	–	A+	0.63	–	–	A+	0.83	A+	0.55	A-	–0.19	A+	0.19
3 *	–	–	B+	1.23	–	–	A+	0.07	A+	0.84	A+	0.41	A-	–0.54
4	–	–	A+	0.88	–	–	A-	–0.17	A+	0.28	A+	0.02	A-	–0.66
5	–	–	A+	0.25	–	–	A+	0.37	A+	0.17	A+	0.24	A+	0.26
6	–	–	A-	–0.03	–	–	B+	1.94	A+	0.46	A+	0.20	A+	0.39
7	–	–	A-	–0.09	–	–	B-	-1.70	A-	–0.52	A-	–0.06	A-	–0.28
8 *	–	–	A+	0.55	–	–	A+	0.47	A+	0.50	A-	–0.26	A-	–0.78
9	–	–	A+	0.20	–	–	A+	0.03	A+	0.11	A-	–0.09	A+	0.15
10	–	–	A-	–0.16	–	–	A+	0.89	A+	0.01	A-	–0.20	A-	–0.10
11 *	–	–	A+	0.07	–	–	B+	1.02	A+	0.37	A+	0.40	A+	0.56
12	–	–	A-	–0.84	–	–	A-	–0.79	A-	–0.84	A-	–0.65	A-	–0.56
13	–	–	A+	0.14	–	–	A+	0.03	A+	0.06	A+	0.03	A-	–0.39
14	–	–	A+	0.54	–	–	A+	0.51	A+	0.42	A+	0.27	A+	0.58
15 *	–	–	A-	–0.27	–	–	A-	–0.21	A-	–0.17	A+	0.16	A-	–0.55
16	–	–	A+	0.49	–	–	A+	0.78	A+	0.63	A-	–0.01	A+	0.22
17	–	–	A+	0.42	–	–	A+	0.18	A+	0.40	A+	0.55	A+	0.79
18	–	–	A-	–0.56	–	–	A+	0.07	A-	–0.39	A+	0.06	A+	0.00
19	–	–	A-	–0.50	–	–	A-	–0.15	A-	–0.29	A-	–0.73	A-	–0.23
20 *	–	–	A-	–0.04	–	–	A+	0.33	A-	–0.03	A-	–0.15	A-	–0.40
21	–	–	A-	–0.56	–	–	A+	0.25	A-	–0.31	A-	–0.08	A+	0.04
22 *	–	–	A-	–0.01	–	–	A+	0.79	A+	0.20	A+	0.03	A-	–0.64
23	–	–	A-	–0.51	–	–	A+	0.11	A-	–0.20	A-	–0.16	A-	–0.04
24	–	–	A-	–0.09	–	–	A-	–0.16	A-	–0.16	A-	–0.16	A+	0.24
25	–	–	A+	0.19	–	–	A-	–0.21	A-	–0.02	A-	–0.24	A-	–0.86
26 *	–	–	A-	–0.12	–	–	A+	0.32	A-	–0.02	A-	–0.54	A-	–0.23
27 *	–	–	B-	-1.15	–	–	B-	-1.42	B-	-1.25	A-	–0.76	A-	–0.64
28	–	–	A-	–0.33	–	–	A-	–0.01	A-	–0.06	A-	–0.14	A-	–0.60
29	–	–	A-	–0.51	–	–	A-	–0.21	A-	–0.36	A-	–0.23	A+	0.02

Item	White/ American Indian †		White/Asian American		White/Pacific Islander†		White/ Filipino		White/ Combined Asian		White/ Hispanic		White/ African American	
	Cat.	MH D-DIF	Cat.	MH D-DIF	Cat.	MH D-DIF	Cat.	MH D-DIF	Cat.	MH D-DIF	Cat.	MH D-DIF	Cat.	MH D-DIF
30	–	–	A+	0.41	–	–	A+	0.34	A+	0.50	A+	0.65	A+	0.44
31	–	–	A-	–0.05	–	–	A+	0.43	A+	0.08	A+	0.18	A+	0.47
32 *	–	–	A-	–0.62	–	–	A-	–0.23	A-	–0.44	A-	–0.51	A-	–0.34
33 *	–	–	A-	–0.36	–	–	A-	–0.91	A-	–0.40	A+	0.03	A-	–0.05
34 *	–	–	A-	–0.59	–	–	A+	0.30	A-	–0.30	A-	–0.22	A-	–0.08
35 *	–	–	A+	0.05	–	–	A+	0.49	A+	0.22	A-	–0.31	A-	–0.46
36 *	–	–	A+	0.03	–	–	A-	–0.70	A-	–0.17	A+	0.30	A+	0.35
37 *	–	–	A+	0.45	–	–	A+	0.52	A+	0.47	A-	–0.49	A+	0.05
38 *	–	–	A+	0.15	–	–	A-	–0.06	A+	0.15	A-	–0.41	A-	–0.29
39	–	–	A-	–0.27	–	–	A+	0.90	A+	0.11	A+	0.11	A+	0.11
40	–	–	A+	0.25	–	–	A+	0.06	A+	0.17	A-	–0.39	A-	–0.25
41 *	–	–	A-	–0.02	–	–	A+	0.12	A-	–0.03	A-	–0.01	A+	0.80
42	–	–	A-	–0.32	–	–	A-	–0.04	A-	–0.27	A-	–0.22	A-	–0.08
43	–	–	A+	0.31	–	–	A+	0.89	A+	0.39	A+	0.10	A-	–0.16
44*	–	–	A+	0.26	–	–	A+	1.25	A+	0.47	A-	–0.25	A-	–0.78
45	–	–	A-	–0.33	–	–	A+	0.39	A-	–0.10	A-	–0.33	A-	–0.68
46	–	–	B-	-1.05	–	–	A+	0.69	A-	–0.44	A-	–0.52	A-	–0.23
47	–	–	A+	0.24	–	–	B+	1.65	A+	0.59	A+	0.25	A-	–0.07
48*	–	–	A+	0.53	–	–	A+	0.30	A+	0.51	A-	–0.29	A-	–0.66
49	–	–	A+	0.43	–	–	A-	–0.16	A+	0.18	A-	–0.44	A-	–0.72
50	–	–	B+	1.41	–	–	A+	1.39	B+	1.50	A-	–0.07	A-	–0.03
51	–	–	A-	–0.39	–	–	A+	0.16	A-	–0.19	A-	–0.09	A-	–0.28
52	–	–	A-	–0.23	–	–	A-	–0.25	A-	–0.20	A-	–0.59	A-	–0.55
53	–	–	A+	0.25	–	–	A+	0.71	A+	0.38	A-	–0.11	A+	0.11
54 *	–	–	A+	0.12	–	–	A-	–0.09	A-	–0.07	A-	–0.55	A-	–0.57
55	–	–	A-	–0.23	–	–	B-	-1.19	A-	–0.51	A-	–0.41	A-	–0.15
56	–	–	A+	0.35	–	–	A+	0.77	A+	0.43	A+	0.22	A-	–0.10
57	–	–	A-	–0.34	–	–	A+	0.51	A-	–0.16	A-	–0.20	A-	–0.11
58	–	–	A+	0.14	–	–	A+	0.06	A+	0.14	A-	–0.18	A-	–0.48
59	–	–	A-	–0.90	–	–	A+	0.03	A-	–0.66	A-	–0.18	A-	–0.33
60	–	–	B+	1.56	–	–	B+	1.03	B+	1.34	A+	0.07	A-	–0.21

* Technology-enhanced item

† Insufficient sample size for DIF analysis

Differential Impact at the Overall Test Level

Table E.19. Effect Sizes for the Comparison of Subgroup Residual Means—Grade Five

Group 1	Group 2	Cohen's *d*
Male	Female	0.05
American Indian	Asian American	–0.04
American Indian	Pacific Islander	0.17
American Indian	Filipino	0.09
American Indian	Hispanic	0.32
American Indian	African American	0.26
American Indian	White	–0.01
American Indian	Two or more races	0.19
Asian American	Pacific Islander	0.20
Asian American	Filipino	0.14
Asian American	Hispanic	0.36
Asian American	African American	0.29
Asian American	White	0.03
Asian American	Two or more races	0.24
Pacific Islander	Filipino	–0.05
Pacific Islander	Hispanic	0.15
Pacific Islander	African American	0.09
Pacific Islander	White	–0.17
Pacific Islander	Two or more races	0.04
Filipino	Hispanic	0.21
Filipino	African American	0.14
Filipino	White	–0.11
Filipino	Two or more races	0.09
Hispanic	African American	–0.06
Hispanic	White	–0.32
Hispanic	Two or more races	–0.10
African American	White	–0.25
African American	Two or more races	–0.04
White	Two or more races	0.21
No special education services	Special education services	0.02
English only	Initially fluent English proficient	0.00
English only	English learner	0.25
English only	Reclassified fluent English proficient	0.05
Initially fluent English proficient	English learner	0.26
Initially fluent English proficient	Reclassified fluent English proficient	0.05
English learner	Reclassified fluent English proficient	–0.21
Not economically disadvantaged	Economically disadvantaged	0.25
Graduate school/Postgraduate training	College graduate	0.02
Graduate school/Postgraduate training	Some college (includes AA degree)	0.09
Graduate school/Postgraduate training	High school graduate	0.24
Graduate school/Postgraduate training	Not a high school graduate	0.31
College graduate	Some college (includes AA degree)	0.07
College graduate	High school graduate	0.23
College graduate	Not a high school graduate	0.30

Group 1	Group 2	Cohen's *d*
Some college (includes AA degree)	High school graduate	0.16
Some college (includes AA degree)	Not a high school graduate	0.23
High school graduate	Not a high school graduate	0.06
Metropolitan	Small town/Rural	–0.11

Table E.20. Effect Sizes for the Comparison of Subgroup Residual Means—Grade Eight

Group 1	Group 2	Cohen's *d*
Male	Female	0.13
American Indian	Asian American	–0.47
American Indian	Pacific Islander	–0.11
American Indian	Filipino	–0.71
American Indian	Hispanic	–0.26
American Indian	African American	–0.26
American Indian	White	–0.65
American Indian	Two or more races	–0.47
Asian American	Pacific Islander	0.37
Asian American	Filipino	–0.15
Asian American	Hispanic	0.25
Asian American	African American	0.24
Asian American	White	–0.15
Asian American	Two or more races	0.09
Pacific Islander	Filipino	–0.61
Pacific Islander	Hispanic	–0.15
Pacific Islander	African American	–0.15
Pacific Islander	White	–0.55
Pacific Islander	Two or more races	–0.36
Filipino	Hispanic	0.42
Filipino	African American	0.45
Filipino	White	0.00
Filipino	Two or more races	0.29
Hispanic	African American	0.01
Hispanic	White	–0.41
Hispanic	Two or more races	–0.16
African American	White	–0.41
African American	Two or more races	–0.18
White	Two or more races	0.25
No special education services	Special education services	0.09
English only	Initially fluent English proficient	0.06
English only	English learner	0.32
English only	Reclassified fluent English proficient	0.14
Initially fluent English proficient	English learner	0.27
Initially fluent English proficient	Reclassified fluent English proficient	0.08
English learner	Reclassified fluent English proficient	–0.18
Not economically disadvantaged	Economically disadvantaged	0.31
Graduate school/Postgraduate training	College graduate	0.08
Graduate school/Postgraduate training	Some college (includes AA degree)	0.28

Group 1	Group 2	Cohen's *d*
Graduate school/Postgraduate training	High school graduate	0.40
Graduate school/Postgraduate training	Not a high school graduate	0.41
College graduate	Some college (includes AA degree)	0.19
College graduate	High school graduate	0.31
College graduate	Not a high school graduate	0.32
Some college (includes AA degree)	High school graduate	0.12
Some college (includes AA degree)	Not a high school graduate	0.13
High school graduate	Not a high school graduate	0.01
Metropolitan	Small town/Rural	–0.12

Table E.21. Effect Sizes for the Comparison of Subgroup Residual Means—Biology

Group 1	Group 2	Cohen's *d*
Male	Female	–0.02
American Indian	Asian American	–0.10
American Indian	Pacific Islander	–0.18
American Indian	Filipino	–0.28
American Indian	Hispanic	–0.08
American Indian	African American	–0.06
American Indian	White	–0.21
American Indian	Two or more races	–0.29
Asian American	Pacific Islander	–0.06
Asian American	Filipino	–0.17
Asian American	Hispanic	0.03
Asian American	African American	0.05
Asian American	White	–0.11
Asian American	Two or more races	–0.16
Pacific Islander	Filipino	–0.11
Pacific Islander	Hispanic	0.09
Pacific Islander	African American	0.12
Pacific Islander	White	–0.05
Pacific Islander	Two or more races	–0.11
Filipino	Hispanic	0.21
Filipino	African American	0.23
Filipino	White	0.05
Filipino	Two or more races	0.01
Hispanic	African American	0.02
Hispanic	White	–0.14
Hispanic	Two or more races	–0.20
African American	White	–0.16
African American	Two or more races	–0.23
White	Two or more races	–0.04
No special education services	Special education services	0.14
English only	Initially fluent English proficient	–0.02
English only	English learner	0.25
English only	Reclassified fluent English proficient	0.00

Group 1	Group 2	Cohen's *d*
Initially fluent English proficient	English learner	0.29
Initially fluent English proficient	Reclassified fluent English proficient	0.02
English learner	Reclassified fluent English proficient	–0.26
Not economically disadvantaged	Economically disadvantaged	0.08
Graduate school/Postgraduate training	College graduate	–0.05
Graduate school/Postgraduate training	Some college (includes AA degree)	0.00
Graduate school/Postgraduate training	High school graduate	0.12
Graduate school/Postgraduate training	Not a high school graduate	0.12
College graduate	Some college (includes AA degree)	0.06
College graduate	High school graduate	0.19
College graduate	Not a high school graduate	0.19
Some college (includes AA degree)	High school graduate	0.12
Some college (includes AA degree)	Not a high school graduate	0.12
High school graduate	Not a high school graduate	0.00
Metropolitan	Small town/Rural	0.43

Factor Analyses

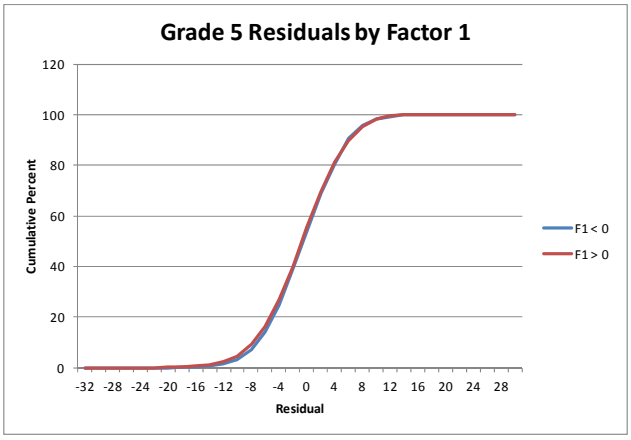


Figure E.2. Cumulative Distributions of Grade Five Science Regression Residuals by Level of Factor 1: Exposure to Computer-related Technologies

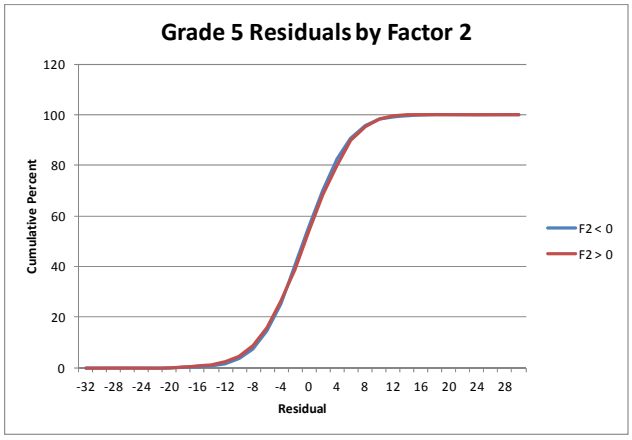


Figure E.3. Grade Five Science Regression Residuals by Level of Factor 2: Exposure to Computer Technologies in a Nonacademic Environment

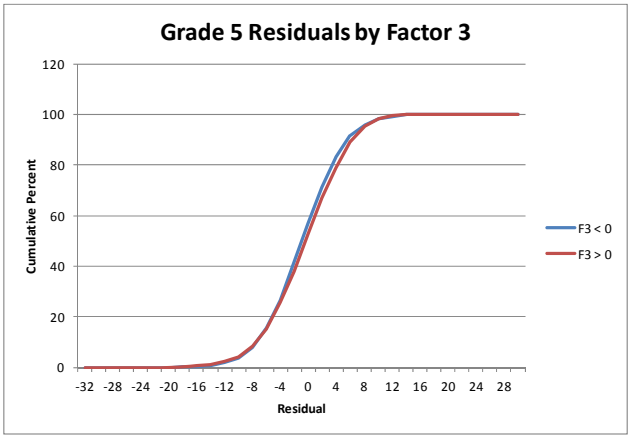


Figure E.4. Cumulative Distributions of Grade Five Science Regression Residuals by Level of Factor 3: Exposure to Computer Technologies in Academics

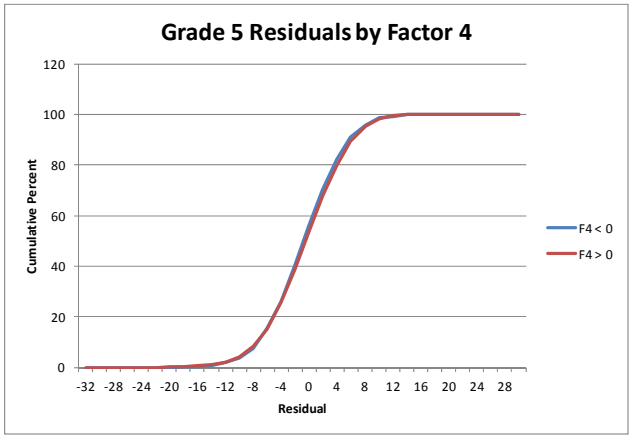


Figure E.5. Cumulative Distributions of Grade Five Science Regression Residuals by Level of Factor 4: Efficacy and Attitude to Computer Technologies

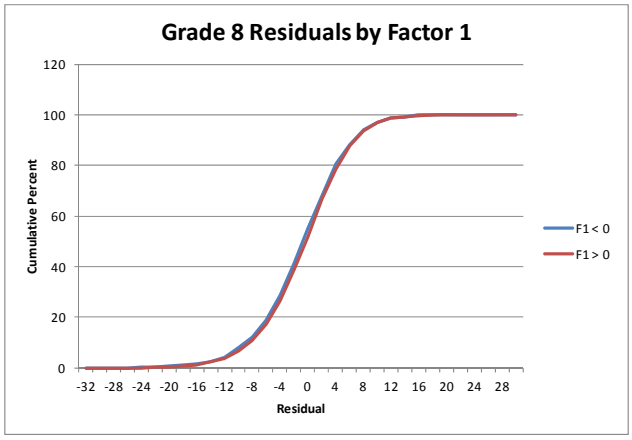


Figure E.6. Grade Eight Science Regression Residuals by Factor 1: Exposure to Computer-related Technologies

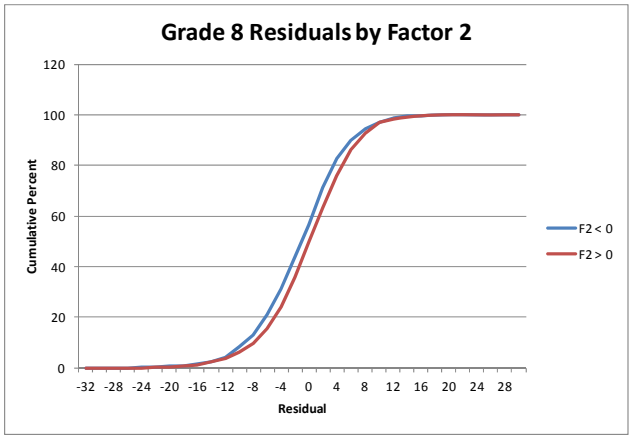


Figure E.7. Cumulative Distributions of Grade Eight Science Regression Residuals by Factor 2: Exposure to Computer Technologies in a Nonacademic Environment

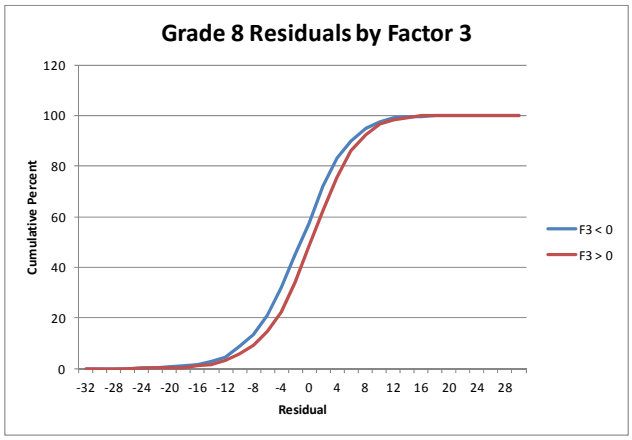


Figure E.8. Cumulative Distributions of Grade Eight Science Regression Residuals by Factor 3: Exposure to Computer Technologies in Academics

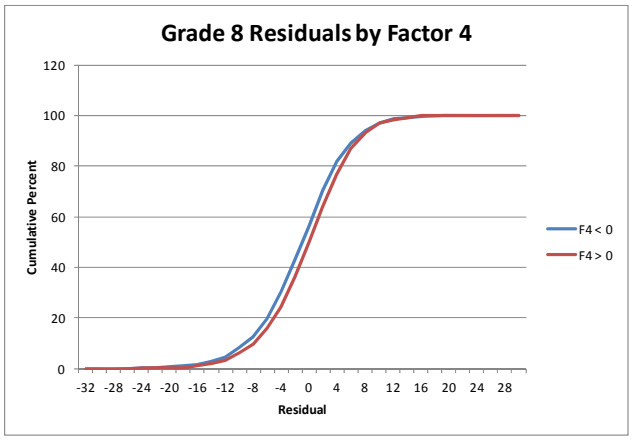


Figure E.9. Cumulative Distributions of Grade Eight Science Regression Residuals by Factor 4: Efficacy and Attitude to Computer Technologies

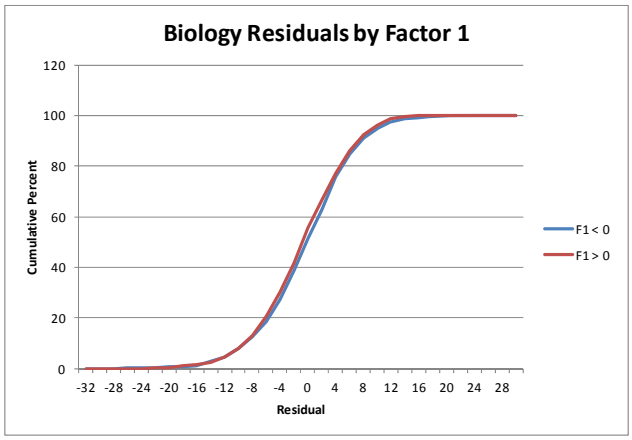


Figure E.10. Cumulative Distributions of Biology Regression Residuals by Factor 1: Exposure to Computer-related Technologies

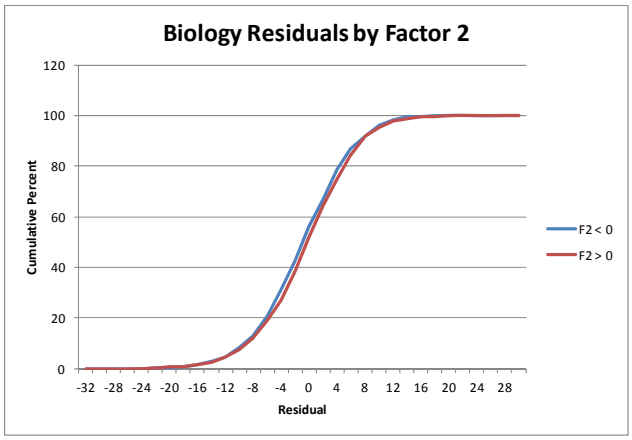


Figure E.11. Cumulative Distributions of Biology Regression Residuals by Factor 2: Exposure to Computer Technologies in a Nonacademic Environment

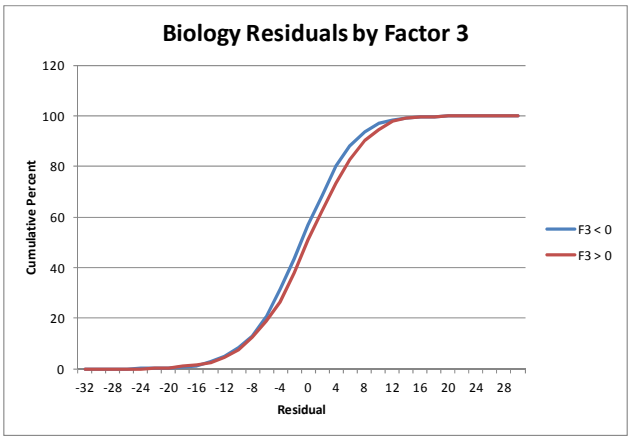


Figure E.12. Cumulative Distributions of Biology Regression Residuals by Factor 3: Exposure to Computer Technologies in Academics

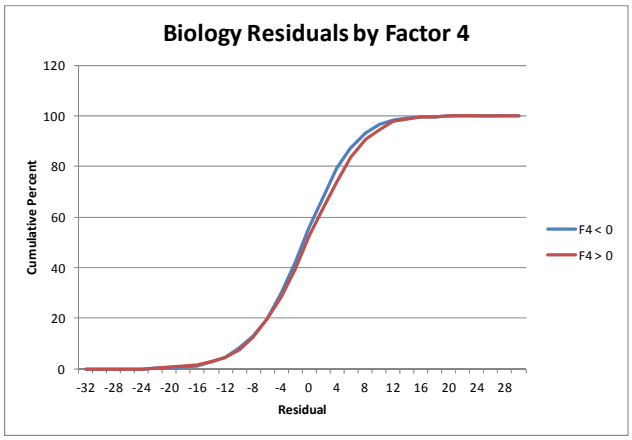


Figure E.13. Cumulative Distributions of Biology Regression Residuals by Factor 4: Efficacy and Attitude to Computer Technologies

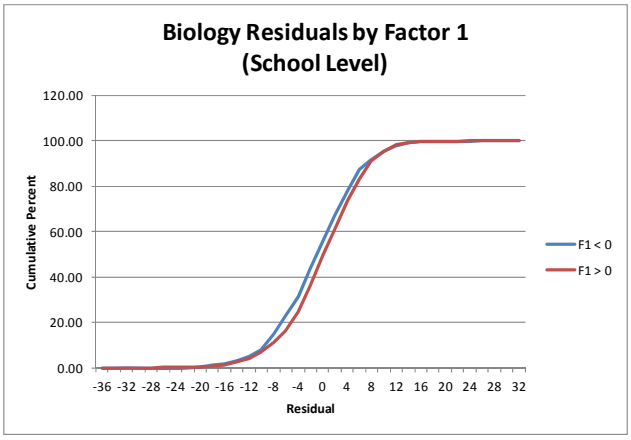


Figure E.14. Cumulative Distributions of Biology Regression Residuals by School Factor 1:Teachers' experience with technology

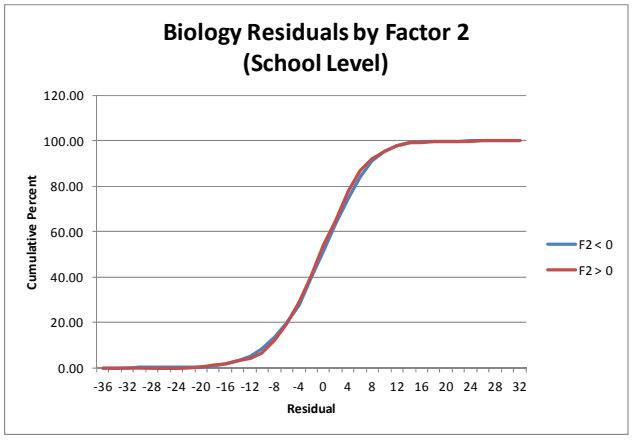


Figure E.15. Cumulative Distributions of Biology Regression Residuals by School Factor 2: School Computer Testing Experience

References

References

Akaike, H. (1987). Factor analysis and AIC. *Psychometrika*, 52, 317–322.

Bridgeman, B. (1992). A comparison of quantitative questions in open-ended and multiple-choice formats. *Journal of Educational Measurement, 29*(3), 253–271.

Brown, T. A. (2006). *Confirmatory factor analysis for applied research*. New York, NY: Guilford Press.

California Department of Education. (2012a). *2012 STAR Computer-based Testing Tryout directions for administration.* Sacramento, CA. Downloaded from https://star.caltesting.org/docs/2012/intro.pdf

California Department of Education. (2012b). *2012 STAR Computer-based Testing Tryout manual.* Sacramento, CA. Downloaded from https://star.caltesting.org/docs/2013/STAR_CBT_Tryout_Manual.pdf

California Department of Education. (2012c). *STAR Computer-based Testing (CBT) Tryout STAR coordinator checklist*. Sacramento, CA. Downloaded from https://star.caltesting.org/docs/2013/STAR_Coordinator_Checklist.pdf

California Department of Education. (2012d). *STAR Computer-based Testing (CBT) Tryout technology coordinator checklist*. Sacramento, CA. Downloaded from https://star.caltesting.org/docs/2013/Technology_Coordinator_Checklist.pdf

California Department of Education. (2012e). *STAR Computer-based Testing (CBT) Tryout test administrator checklist*. Sacramento, CA. Downloaded from https://star.caltesting.org/docs/2013/Test_Administrator_Checklist.pdf

California Department of Education. (2012f). *STAR Test security affidavit for test examiners, proctors, translators, scribes, and any other person having access to the STAR Computer-based Testing (CBT) Tryout for Science*. Sacramento, CA. Downloaded from https://star.caltesting.org/docs/2013/STAR_CBT_Tryout_Security_affidavit.pdf

Cattell, R. B. (1966). The scree test for the number of factors. *Multivariate Behavioral Research*, 1(2), 245–276.

Cohen, J. A. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Cohen, J. A. (1992). Power primer. *Psychological Bulletin, 112,* 55–159.

Comrey, A. L. (1988). Factor-analytic methods of scale development in personality and clinical psychology. *Journal of Consulting and Clinical Psychology*, 56, 754–761.

Cronbach, L. J. (1951). Coefficient alpha and the independent structure of tests. *Psychometrika, 16*, 292-334.

Educational Testing Service. (2012a). *Configure the TDE to connect to the LCS (Mac).* [Video file]. Retrieved from https://vimeo.com/49916202

Educational Testing Service. (2012b). *Configure the TDE to connect to the LCS (PC).* [Video file]. Retrieved from http://vimeo.com/49289521

Educational Testing Service. (2012c). *Download and install the LCS (PC).* [Video file]. Retrieved from http://vimeo.com/49177058

Educational Testing Service. (2012d). *Download and install the LCS (Mac).* [Video file]. Retrieved from https://vimeo.com/49845168

Educational Testing Service. (2012e). *Download and install the TDE (Mac).* [Video file]. Retrieved from https://vimeo.com/49863254

Educational Testing Service. (2012f). *Download and install the TDE (PC).* [Video file]. Retrieved from http://vimeo.com/49254240

Educational Testing Service. (2012g, September 14). *Preparing for and administering the STAR CBT Tryout (for test administrators and STAR coordinators).* [Video file]. Retrieved from http://www.startest.org/CBT-admin-webcast.2012.html

Educational Testing Service. (2012h). *Readiness check - part 1 (accessing the LCS monitoring page) (Mac).* [Video file]. Retrieved from https://vimeo.com/49925475

Educational Testing Service. (2012i). *Readiness check - part 1: Accessing the LCS monitoring page (PC).* [Video file]. Retrieved from http://vimeo.com/49272356

Educational Testing Service. (2012j). *Readiness check - part 2 (monitoring a practice test) (Mac).* [Video file]. Retrieved from https://vimeo.com/49938028

Educational Testing Service. (2012k). *Readiness check - part 2: Monitoring a practice test (PC).* [Video file]. Retrieved from http://vimeo.com/49267525

Educational Testing Service. (2012l). *Running connectivity in the LCS Manager (Mac).* [Video file]. Retrieved from https://vimeo.com/49854316

Educational Testing Service. (2012m). *Running connectivity in the LCS Manager (PC).* [Video file]. Retrieved from http://vimeo.com/49290293

Educational Testing Service. (2012n, September 14). *Setting up your computer systems for the STAR CBT Tryout (for technology coordinators).* [Video file]. Retrieved from http://www.startest.org/CBT-tech-setup-webcast.2012.html

Educational Testing Service. (2012o). *Student tutorials* [Online tutorial]. Princeton, NJ: Author. Retrieved from https://star.caltesting.org/tutorials/index.html

Educational Testing Service. (2012p). *System requirements.* Retrieved from http://www.startest.org/pdfs/CBT_tryout.system_req.2012.pdf

Educational Testing Service. (2012q). *Troubleshooting (PC).* [Video file]. Retrieved from http://vimeo.com/49243668

Enders, C. K., & Tofighi, D. (2008). The impact of misspecifying class-specific residual variances in growth mixture models. *Structural Equation Modeling: A Multidisciplinary Journal*, 15, 75–95.

Fabrigar, L. R., Wegener, D. T., MacCallum, R. C., & Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods*, 4, 272–299.

Gaffney, T., & Perryman, C. (2009). *A longitudinal look at the factor structure of educational achievement tests*. Paper presented at the annual meeting of the Psychometric Society, Cambridge, England.

Holland, P. W. & Thayer, D. T. (1985). *An alternative definition of the ETS delta scale of item difficulty* (Research Report 85–43). Princeton, NJ: Educational Testing Service.

Hu, L. T., & 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, 1–55.

Mantel, N. & Haenszel, W. (1959). Statistical aspects of the analyses of data from retrospective studies of disease. *Journal of the National Cancer Institute*, 22, 719–48.

Leacock, C., & Chodorow, M. (2003). C-rater: Scoring of short-answer questions. *Computers and the Humanities, 37(4),* 389–405.

Muthen, L. K., & Muthen, B. O. (1998). *Mplus user's guide* (5th ed.). Los Angeles, CA: Los Angeles, CA: Muthen & Muthen.

SAS Institute Inc. (2009). *Base SAS® 9.2 procedures guide*. Cary, NC: SAS Institute Inc.

Schwarz, G. (1978). Estimating the dimension of a model. *Annals of Statistics*, 6, 461-464.

Williamson, D. M., Xi, X., & Breyer, F. J. (2012). A framework for evaluation and use of automated scoring. *Educational Measurement: Issues and Practice, 31*(1), 2–13.