



RIDE Rhode Island
Department
of Education

2019 RICAS Technical Report

Prepared by Cognia and the
Rhode Island Department of Education



Table of Contents

CHAPTER 1 OVERVIEW.....	5
1.1 PURPOSE OF THIS REPORT.....	5
1.2 THE RHODE ISLAND COMPREHENSIVE ASSESSMENT SYSTEM.....	5
1.2.1 Appropriateness of Using Massachusetts Standards.....	6
1.2.2 Content Standards.....	6
1.2.3 Performance Standards.....	6
1.3 ORGANIZATION OF THIS REPORT.....	7
CHAPTER 2 TEST ADMINISTRATION.....	9
2.1 TEST ADMINISTRATION SCHEDULE.....	9
2.2 SECURITY REQUIREMENTS.....	9
2.3 PARTICIPATION REQUIREMENTS.....	10
2.4 STUDENTS NOT TESTED ON STANDARD TESTS.....	10
2.4.1 Special Edition Test Forms.....	10
2.5 ADMINISTRATION PROCEDURES.....	11
CHAPTER 3 SCORING.....	12
3.1 INTERRATER CONSISTENCY.....	12
CHAPTER 4 REPORTING.....	14
4.1 REPORTING OF RESULTS.....	14
4.2 PARENT/GUARDIAN REPORT.....	14
4.3 STUDENT RESULTS LABELS.....	15
4.4 DECISION RULES.....	15
4.5 QUALITY ASSURANCE.....	15
CHAPTER 5 PSYCHOMETRIC QUALITY.....	16
5.1 CLASSICAL ITEM ANALYSES.....	16
5.1.1 Classical Difficulty and Discrimination Indices.....	16
5.1.2 Differential Item Functioning.....	18
5.1.3 Dimensionality Analysis.....	19
5.2 RICAS RELIABILITY.....	20
5.2.1 Reliability and Standard Errors of Measurement.....	20
5.2.2 Reporting Subcategory Reliability.....	20
5.2.3 Subgroup Reliability.....	21
5.2.4 Decision Accuracy and Consistency Results.....	21
REFERENCES.....	25
APPENDICES.....	26

Appendix A	Participation Rates
Appendix B	Accommodations
Appendix C	Interrater Consistency
Appendix D	Achievement Level Distributions

Appendix E	Item-Level Classical Statistics
Appendix F	Score Distributions
Appendix G	Differential Item Functioning Results
Appendix H	Reliability

List of Tables

Table 1-1. Relationship between 2019 RICAS and MCAS Tests on Critical Test Components	7
Table 2-1. Grades 3–8 ELA and Mathematics Test Administration Schedule	9
Table 3-1. Summary of Interrater Consistency Statistics Organized across Items by Content Area and Grade—ELA	12
Table 3-2. Summary of Interrater Consistency Statistics Organized across Items by Content Area and Grade—Mathematics.....	13
Table 5-1. Summary of Item Difficulty and Discrimination Statistics by Content Area and Grade	17
Table 5-2. Multidimensionality Effect Sizes by Grade and Content Area	19
Table 5-3. Raw Score Descriptive Statistics, Cronbach’s Alpha, and SEMs by Content Area and Grade	20
Table 5-4. Summary of Decision Accuracy (and Consistency) Results by Content Area and Grade—Overall and Conditional on Achievement Level	22
Table 5-5. Summary of Decision Accuracy (and Consistency) Results by Content Area and Grade—Conditional on Cutpoint.....	23

Chapter 1 OVERVIEW

1.1 PURPOSE OF THIS REPORT

The purpose of this 2019 Rhode Island Comprehensive Assessment System (RICAS) Technical Report is to document the technical quality and characteristics of the 2019 RICAS English Language Arts and Mathematics tests in grades 3–8, in order to present evidence of the validity, reliability, and fairness of the use of the tests as part of the Rhode Island state assessment program.

Because the tests administered in RICAS are the MCAS English Language Arts and Mathematics tests, most of the information related to their technical quality is provided in the MCAS Technical Reports produced by the Massachusetts Department of Elementary and Secondary Education (MA DESE). MCAS Technical Reports are available on the MA DESE website: doe.mass.edu/mcas/tech/?section=techreports.

This report contains information specific to the administration of the tests in Rhode Island. It is intended to accompany the information contained in the MCAS Technical Report, document any differences in the assessment policies and procedures between Rhode Island and Massachusetts, and provide additional background information about the RICAS program.

The information contained in this report, in conjunction with the technical documentation prepared by Massachusetts, demonstrates that the grades 3–8 MCAS English Language Arts and Mathematics tests are technically sound, function well for students in Rhode Island, and are appropriate instruments to assess the performance of Rhode Island students on the state's content standards.

This report is primarily intended for experts in psychometrics and educational measurement. It assumes a working knowledge of measurement concepts, such as reliability and validity, as well as statistical concepts of correlation and central tendency. For some sections, the reader is presumed to have basic familiarity with advanced topics in measurement and statistics, such as item response theory (IRT) and factor analysis.

1.2 THE RHODE ISLAND COMPREHENSIVE ASSESSMENT SYSTEM

The RICAS is Rhode Island's state assessment program in English language arts and mathematics at grades 3–8, designed to meet the federal requirements of the Every Student Succeeds Act (ESSA). In addition to fulfilling ESSA assessment requirements, the specific purposes of the RICAS tests are (1) to provide information to parents/guardians and students on Rhode Island student achievement on the state's English language arts and mathematics content standards, (2) to provide information to support program evaluation and improvement at the school and district level, and (3) to provide academic achievement and growth information used as part of the state's school accountability program to inform parents/guardians and the public about the performance of Rhode Island schools.

Beginning in the 2017–2018 school year, the Rhode Island Department of Education (RIDE) adopted the MCAS English Language Arts and Mathematics tests as its state assessments in English language arts and mathematics at grades 3–8. The tests are administered in Rhode Island under a licensing agreement with MA DESE and labeled RICAS for their use in Rhode Island. The use of the MCAS tests at grades 3–8 is part of Rhode Island's transition from the use of the Partnership for the Assessment of Readiness for College and Careers (PARCC) tests at grades 3–8 and high school as its state assessments. In high school, the PARCC tests have been replaced by the SAT.

The adoption of the MCAS tests reflects a continuation of Rhode Island's policy to partner with other states to offer a high-quality state assessment. With the increased assessment requirements of the No Child Left Behind Act in 2001, RIDE determined that it would not be feasible to develop and sustain a high-quality assessment program on its own. From 2003–2014, Rhode Island partnered with New Hampshire, Vermont, and Maine in the New England Common Assessment Program (NECAP). With the adoption of the Common Core State Standards (CCSS) and the creation of national assessment consortia, Rhode Island joined PARCC, administering the PARCC tests from 2015–2017.

As Massachusetts and other states left the PARCC consortium, it was no longer clear that PARCC would be able to offer long-term stability in assessment to support the state’s improvement efforts. MCAS, in contrast, has been regarded as a model for high-quality and stable state assessment since its inception in 1998. In 2017, Massachusetts developed MCAS tests to fully align with college- and career-ready content standards and established rigorous performance standards consistent with those established by PARCC. With the updated tests and performance standards in place, Rhode Island began administration of the Massachusetts tests in spring 2018.

1.2.1 Appropriateness of Using Massachusetts Standards

Before adopting the MCAS tests as its state assessment, it was necessary to determine the appropriateness of the Massachusetts content and performance standards for use in Rhode Island. To meet ESSA requirements and provide valid and useful information to Rhode Island parents/guardians, students, and schools, the state assessments must be aligned to the state’s content standards. In addition, to support the state’s commitment to ensure that Rhode Island’s educational system holds high expectations for all students and that Rhode Island graduates are well prepared for postsecondary education, work, and life, the state must establish rigorous performance standards that signal whether students are on track for success in high school and college and career readiness as they progress through elementary and middle school.

1.2.2 Content Standards

In 2010, Rhode Island adopted the CCSS as its state content standards in English language arts and mathematics. In July 2010, the Massachusetts Board of Elementary and Secondary Education also adopted the CCSS in English language arts and mathematics as the core of its PK–12 content standards.

In March 2011, Massachusetts adopted revised *Curriculum Frameworks* in English language arts and mathematics, which are the state’s academic content standards. As described at the time by Mitchell Chester, Massachusetts Commissioner of Elementary and Secondary Education, the 2011 *Curriculum Frameworks*, “merges the *Common Core State Standards for Mathematics* with additional Massachusetts standards and other features.” In English language arts, the elements unique to Massachusetts were described as including standards for pre-kindergartners, expansions of the *Common Core*’s glossary and bibliography, and two sections that suggest appropriate classic and contemporary authors for different grade-level ranges. In mathematics, the elements unique to Massachusetts were described as including standards for pre-kindergartners, Guiding Principles for mathematics programs, expansions of the *Common Core*’s glossary and bibliography, and an adaptation of the CCSS high school model courses.

The CCSS remain the core to which the MCAS is aligned. This is particularly true at grades 3–8, in which the MCAS tests are administered in Rhode Island. To support Rhode Island teachers’ understanding of the correspondence between the CCSS and the Massachusetts *Curriculum Frameworks*, RIDE has produced grade-by-grade guides for teachers that articulate the alignment between the CCSS standards and the RICAS tests, documenting any differences between individual CCSS standards and the standards to which the assessments are aligned.

- In English language arts, these Assessment Tables are available on the RIDE website at <http://ride.ri.gov/InstructionAssessment/Assessment/RICASAssessments.aspx#39551541-test-design-english-language-arts-information>.
- In mathematics, these guides are presented as Assessment Tables and Achievement Level Descriptors and are available on the RIDE website at <http://www.ride.ri.gov/InstructionAssessment/Assessment/RICASAssessments.aspx#39551515-test-design-mathematics-information>.

1.2.3 Performance Standards

In addition to the alignment of the tests to Rhode Island’s academic content standards, for the MCAS tests to be appropriate for Rhode Island it was essential that the performance standards established for those tests were

consistent with the rigorous performance standards that Rhode Island adopted when it began administering the PARCC tests in 2015.

Massachusetts conducted standard setting activities in August 2017 to establish achievement level cut scores on the new MCAS tests. RIDE staff and technical advisors observed those standard setting procedures and analyzed the results of the standard setting process. Although results of the new tests are reported in terms of four achievement levels rather than the five levels used to report PARCC results, analyses indicate that the MCAS performance standards are consistent with and as rigorous as the PARCC performance standards previously used in Rhode Island.

Across all grade levels 3–8, results from Rhode Island and Massachusetts suggest that performance at the Meeting Expectations level on the MCAS tests (level 3) is roughly equivalent to performance at the Met Expectations level on the PARCC tests (level 4).

1.3 ORGANIZATION OF THIS REPORT

This report provides information regarding the spring 2019 administration of the 2019 RICAS tests in English language arts and mathematics, including a description and results of analyses conducted to provide evidence of the technical quality and characteristics of those tests.

The RICAS tests were administered, scored, and processed by Cognia, the state’s assessment contractor for the RICAS tests. Cognia is also the Massachusetts assessment contractor for the MCAS tests. Unless noted in this report, all processes and procedures used in administering, processing, scoring, and reporting of the results of the spring 2019 RICAS tests were identical to the corresponding procedures used by Cognia for the MCAS tests. Table 1-1 provides a summary of the relationship between key aspects of the RICAS and MCAS testing programs.

Table 1-1. Relationship between 2019 RICAS and MCAS Tests on Critical Test Components

<i>Test Component</i>	<i>RICAS and MCAS</i>
Test Content	Identical
Test Design	Identical
Test Administration	
• Administration Procedures	Identical
• Mode of Administration	Identical
• Administration Platform	RI offers Spanish language form in mathematics Identical
Scoring	
• Machine-scored items	Identical
• Hand-scored items	Identical
Psychometric Quality	Identical
Reporting	
• Scaled Scores	Identical
• Achievement Levels	Identical

Cognia conducted all the analyses described in this report. The analyses described and presented here are consistent with the types of analyses conducted for the MCAS tests. All analyses are based only on Rhode Island students.

The specific analyses included in this report were identified by the Rhode Island Technical Advisory Committee as necessary and useful to provide evidence of the validity, reliability, and fairness of the use of the MCAS tests as the Rhode Island state assessments in English language arts and mathematics in grades 3–8.

This information includes the following:

- Chapter 2: Test Administration – information related to test administration policies and procedures, including protocols to monitor test security
- Chapter 3: Scoring – information on hand-scoring procedures for short-answer, constructed-response, and essay items, including information on the level of interrater agreement among raters
- Chapter 4: Reporting – detailed information on the type of student-level test scores reported to parents/guardians and a description of the quality assurance procedures used to ensure the accuracy of the reporting of those results
- Chapter 5: Psychometric Quality – a description of and summary results from the set of analyses conducted with Rhode Island students to demonstrate the technical quality and characteristics of the tests (Statistics provided include Classical Item Statistics; Differential Item Functioning; Reliability, including subgroup reliability; and Decision Consistency/Accuracy.)

Additionally, a set of appendices is provided, containing the following information:

- Appendix A – Participation Rates
- Appendix B – Accommodations
- Appendix C – Achievement Level Distributions
- Appendix D – Item-Level Classical Statistics
- Appendix E – Score Distributions
- Appendix F – Differential Item Functioning Results
- Appendix G – Reliability

Chapter 2 TEST ADMINISTRATION

2.1 TEST ADMINISTRATION SCHEDULE

The standard grades 3–8 RICAS tests were administered in two modes, computer-based and paper-based during two overlapping periods in spring 2019, as shown in Table 2-1:

Table 2-1. Grades 3–8 ELA and Mathematics Test Administration Schedule

<i>Content Area</i>	<i>Complete the Student Registration/ Personal Needs Profile (SR/PNP) Process</i>	<i>Receive Test Administration Materials</i>	<i>Test Administration Windows</i>	<i>Deadline for Return of Materials to Contractor (for PBT Only)</i>
ELA and Mathematics	January 28 – March 15, 2019	March 18–20, 2019	April 1 – May 3, 2019 for ELA April 2 – May 24, 2019 for Mathematics	May 7, 2019 for ELA May 29, 2019 for Mathematics

2.2 SECURITY REQUIREMENTS

Principals were responsible for ensuring that all test administrators complied with the requirements and instructions contained in the *Test Coordinator’s Manual* and *Test Administrator’s Manuals*. In addition, other administrators, educators, and staff within the school were responsible for complying with the same requirements. Schools and school staff who violated the test security requirements were subject to numerous possible sanctions and penalties, including employment consequences, delays in reporting of test results, the invalidation of test results, the removal of school personnel from future RICAS administrations, and possible licensure consequences for licensed educators.

If test content is breached, quick identification and resolution of the breach are critical to the integrity of a testing program. In addition to reports of breaches in the field, the RICAS program used the Pearson propriety web monitoring tool to perform web monitoring. The Pearson web monitoring system leverages technology tools and human expertise to identify, prioritize, and monitor sites where sensitive test information may be disclosed. The following strategies were used:

- systematically patrolled the Internet, websites, blogs, discussion forums, video archives, social media, document archives, brain dumps, auction sites, and media outlets
- identified and verified threats to RICAS test security and notified Pearson (who notified RIDE and Cognia, as required)
- worked systematically through the steps necessary to have infringing content removed, if a threat was verified
- provided summary reporting that included overall and specific threat analysis

Full security requirements, including details about responsibilities of principals and test administrators, examples of testing irregularities, guidance for establishing and following a document tracking system, and lists of approved and unapproved resource materials, can be found in the *Spring 2019 Test Coordinator’s Manual (TCM), Grades 3–8* and the *2019 Test Administrator’s Manuals (TAMs)*. In spring 2019, there was one TAM for grades 3–8 computer-based tests, one TAM for grades 3-8 paper-based tests. The primary delivery mode was computer-based, with paper-based delivery as accommodation only for students with disabilities.

2.3 PARTICIPATION REQUIREMENTS

Students in grades 3–8 are expected to participate in RICAS tests for the grade in which they are enrolled and reported to RIDE through the enrollment census.

Participation requirements and guidelines for English learner (EL) students and students with significant disabilities are provided in the sections that follow, the participation rates are presented in Appendix A.

See Part III of the *Test Coordinator’s Manual* for information about scheduling test administration, including make-up sessions for students who are absent on the day of testing.

2.4 STUDENTS NOT TESTED ON STANDARD TESTS

A very small number of students educated with Rhode Island public funds were not required to take the standard RICAS tests. These students were strictly limited to the following categories:

- First-year EL students who enrolled in U.S. schools after April 1, 2018, for whom ELA testing is not required. (First-year EL students must participate in RICAS or Dynamic Learning Maps (DLM) mathematics tests.) See the *RICAS Accessibility and Accommodations Manual, 2019* for details on how EL students participate in spring 2019 RICAS.
- Students with significant cognitive disabilities who are eligible for the alternate assessment, The Dynamic Learning Maps Assessment. For more information, refer to the Dynamic Learning Maps (DLM) Alternate Assessments page of the RIDE website:
<https://www.ride.ri.gov/InstructionAssessment/Assessment/DLMAssessments.aspx>
- Rare and unique situations in which a student is unable to participate in statewide assessments due to a documented, significant, and incapacitating emergency that extends across the entire (or remaining) test window.

More details about test administration policies and participation requirements for non-disabled students, for students with disabilities, for EL students, and for students educated in alternate settings can be found in the *Test Coordinator’s Manual*. Data concerning the number of students tested with accommodations is available in Appendix B.

2.4.1 Special Edition Test Forms

Spanish-Speaking Students

Spanish editions of the spring grades 3–8 mathematics test were available to any EL student with a low level of English proficiency who was receiving or had received mathematics instruction in Spanish. The Spanish edition of the grades 3–8 mathematics test contained all common and matrix items found in Form 1 of the operational test.

Cognia employed two independent translators to complete the translation of the grades 3–8 mathematics test to Spanish. The translation process was as follows:

- A set of translation rules or parameters was generated, taking the following into consideration: vocabulary, usage, and consistency over the years. These rules were provided to both translators.
- The first translator translated from English to Spanish. The second translator proofed the work of the first translator. Discrepancies between the two translations were resolved by the first translator.
- The Publishing Department reviewed the graphics in Spanish.

- The script that the teacher read when administering the test was also translated into Spanish and was included as Appendix A of the *Test Administrator's Manual*.

The Spanish editions of the grades 3–8 mathematics test were available in both paper and online formats. Human read aloud in Spanish was also available to students.

2.5 ADMINISTRATION PROCEDURES

It is the test coordinator's responsibility to coordinate the school's RICAS test administration. This coordination responsibility includes the following:

- understanding and enforcing the test security requirements and test administration protocols
- ensuring that students participate in testing according to the requirements in Chapter 2 of this report
- coordinating the school's test administration schedule and ensuring that tests are scheduled during the prescribed testing window, and in the prescribed order
- ensuring that accommodations are properly administered and that transcriptions, if required for any accommodation, are properly completed
- completing the Principal's Certification of Proper Test Administration (PCPA) and ensuring the accuracy of information provided on the form
- providing RIDE with the school's correct contact information

More details about test administration procedures, including ordering test materials, scheduling test administration, designating and training qualified test administrators, identifying testing spaces, meeting with students, providing accurate student information, and accounting for and returning test materials, can be found in the *Test Coordinator's Manual*.

The RICAS program is supported by the RICAS Service Center, which includes a toll-free telephone line and email answered by staff members who provide support to schools and districts. The RICAS Service Center operates weekdays from 7:00 a.m. to 5:00 p.m. (Eastern Time), Monday through Friday.

Chapter 3 SCORING

3.1 INTERRATER CONSISTENCY

Interrater consistency statistics are the result of the processes implemented to ensure valid and reliable hand-scoring of items and, as such, provide evidence of scoring stability. Double-blind scoring was one of the processes used to monitor the quality of the hand-scoring of student responses for constructed-response items. For student constructed-response questions in grades 3–8, 10% were randomly selected and scored independently by two different scorers. Results of the double-blind scoring were used during the scoring process to identify scorers who required retraining or other intervention, and they are presented here as evidence of scoring consistency on the RICAS tests.

A third score was required for any score category in which there was not an exact agreement between scorer one and scorer two. A third score was also required as a confirmation score when either scorer one and/or scorer two provided a score of M for Demonstration of Skills and Concepts and Independence or a score of 1 for Level of Complexity.

Summaries of the interrater consistency results are presented in Tables 3-1 for ELA and 3-2 for mathematics by grade. The tables show the number of score categories, the number of included scores, the percent exact agreement, the percent adjacent agreement, the correlation between the first two sets of scores, and the percent of responses that required a third score. Interrater consistency data is available at the item level in Appendix C.

**Table 3-1. Summary of Interrater Consistency Statistics
Organized across Items by Content Area and Grade—ELA**

Content Area	Grade	Number of		Percent			Correlation	Percentage of Third Scores	Kappa
		Score Categories	Included Scores	Exact	Adjacent	Exact +Adjacent			
ELA	3	4	2,939	72.58	26.71	99.29	0.80	1.33	0.671
		5	1,953	72.86	26.11	98.97	0.81	1.74	0.681
	4	4	3,042	76.00	23.44	99.44	0.83	0.99	0.713
		5	2,014	69.27	30.09	99.36	0.78	1.29	0.635
	5	4	3,134	69.50	29.26	98.76	0.78	3.00	0.665
		5	3,134	67.01	30.73	97.74	0.81	3.00	0.677
	6	4	3,117	70.90	28.33	99.23	0.83	1.80	0.716
		6	3,117	69.04	29.68	98.72	0.85	1.80	0.715
	7	4	3,065	74.42	25.42	99.84	0.87	0.55	0.758
		6	3,065	72.27	27.31	99.58	0.89	0.55	0.748
	8	4	3,015	77.48	22.22	99.70	0.89	1.72	0.799
		6	3,015	74.26	24.25	98.51	0.90	1.72	0.793

**Table 3-2. Summary of Interrater Consistency Statistics
Organized across Items by Content Area and Grade—Mathematics**

<i>Content Area</i>	<i>Grade</i>	<i>Number of</i>		<i>Percent</i>			<i>Correlation</i>	<i>Percentage of Third Scores</i>	<i>Kappa</i>
		<i>Score Categories</i>	<i>Included Scores</i>	<i>Exact</i>	<i>Adjacent</i>	<i>Exact +Adjacent</i>			
Mathematics	3	4	4,124	93.23	6.35	99.58	0.97	0.41	0.804
	4	5	3,965	87.52	11.40	98.92	0.95	1.08	0.905
	5	5	4,202	87.15	11.92	99.07	0.96	0.93	0.897
	6	5	4,495	89.88	9.63	99.51	0.97	0.47	0.902
	7	5	4,102	89.86	9.51	99.37	0.97	0.63	0.897
	8	5	3,010	88.74	11.26	100.00	0.97	0.00	0.956

Chapter 4 REPORTING

4.1 REPORTING OF RESULTS

Results on the RICAS were reported in terms of achievement levels that describe student achievement in relation to established state standards. There are four achievement levels for ELA and mathematics for students in grades 3–8: *Not Meeting Expectations*, *Partially Meeting Expectations*, *Meeting Expectations*, and *Exceeding Expectations*. Students were given a separate achievement-level classification in each content area. Reports are generated at the student level. The achievement level distributions are provided in Appendix D.

Parent/Guardian Reports and student results labels are the only printed reports; they were mailed to districts for distribution to parents/guardians and schools.

4.2 PARENT/GUARDIAN REPORT

The Parent/Guardian Report was generated for each student eligible to take the RICAS tests. The report is a stand-alone single page (11" x 17") color report that is folded. Two full-color copies of each student's report were printed: one for the parent/guardian and one for the school's records. The report is designed to present parents/guardians with a detailed summary of their child's RICAS performance and to enable comparisons with other students at the school, district, and state levels.

The front cover of the Parent/Guardian Report provides student identification information, including student name, grade, date of birth, ID (SASID), school name, and district name. The cover also presents general information about the test, website information for parent/guardian resources, and a summary of the student's results for each content area. This summary provides important information for each content area at a glance, including the student's achievement level, scaled score, range of scores, and growth percentile.

The inside portion of the report contains the achievement level, scaled score, and standard error of the scaled score for each content area tested. If the student does not receive a scaled score, the reason is displayed after "Your Child's Achievement Level." Each achievement level has its own distinct color, and that color is used throughout the report to highlight important report elements based on the student's achievement level and score. These report elements include the student's earned achievement level, scaled score, the visual scale's achievement-level title and achievement-level cut scores, and the comparison of the student's scaled score to the average scaled score at the student's school, district, and the state levels. If the student received a score last year, their earned scaled score from last year's test is also displayed along with the current year scaled score for each content area tested. The previous year's scaled score is displayed in the color corresponding to the achievement level earned that year.

A student growth percentile (SGP) for each content area tested is displayed with a comparison to the average SGP for the student's school and district. The average state level SGP is displayed at the 50th percentile. An SGP describes the student's learning over time compared to his or her academic peers (peers are other students with similar scores on previous state tests).

The student's performance in each content area's reporting categories is also displayed using pictographs and text that indicates the points earned by the student versus the total points possible in that reporting category. For each reporting category, the average number of points earned by students scoring close to 500 is also displayed for comparison purposes. The student's performance on individual test questions is reported at the bottom of the results page in a simplified item response grid. The grid indicates the points earned and points possible for each test question. Essay questions are indicated on the grid. A link to an external resource is also provided for parents/guardians who wish to review test question descriptions on the department's website.

For the first time, in 2019, Parent/Guardian Reports were also available online through Pearson Access Next (PAN). PDFs were run by grade and school and posted online for school, district, and state access. The Parent/Guardian Reports available in PAN are identical to the reports that are printed and shipped.

4.3 STUDENT RESULTS LABELS

The other report that is produced for each student is the *Student Results Label*. The labels are printed and shipped to districts. Each label contains the following information for a student: the student's name, their SASID, grade, date of birth, test date, school code, school name, district name. The student's results for each subject is also reported. Their earned achievement level and scaled score is provided for each subject tested. If the student does not test in one of the subjects, the not tested reason appears on the label. Files are organized by grade, district, and school. Labels are sorted by last name then first name.

4.4 DECISION RULES

To ensure that RICAS results are processed and reported accurately, a document specifying decision rules is prepared before reporting results. The decision rules are adhered to in the processing and analyses of the RICAS test data and in preparation of the reporting results. These rules specify which, if any, student data need to be excluded from school-, district-, and state-level summary computations. At an individual student level, the decision rules document describes how any special cases should be treated for reporting purposes.

4.5 QUALITY ASSURANCE

Quality assurance measures are implemented throughout the process of analysis and reporting at Cognia. The data processors and data analysts perform routine quality-control checks of their computer programs. When data are handed off to different units within the data team, the sending unit verifies that the data are accurate before handoff. Additionally, when a unit receives a data set, the first step is to verify the accuracy of the data. Once new report designs were approved by RIDE, reports were run using demonstration data to test the application of the decision rules. The populated reports were then approved by RIDE.

Another type of quality assurance measure used at Cognia is parallel processing. One data analyst is responsible for writing all programs required to populate the student-level and aggregate reporting tables for the administration. Each reporting table is assigned to a second data analyst who uses the decision rules to independently program the reporting table. The production and quality-assurance tables are compared; when there is 100% agreement, the tables are released for report generation.

The third aspect of quality control involves procedures to check the accuracy of reported data. Using a sample of schools and districts, the quality assurance group verifies that the reported information is correct. The selection of sample schools and districts for this purpose is very specific because it can affect the success of the quality-control efforts. There are two sets of samples selected that may not be mutually exclusive. The first set includes samples that satisfy all the following criteria:

- one-school district
- two-school district
- multi-school district
- private school
- special school (e.g., a charter school)
- small school that does not have enough students to report aggregations
- school with excluded (not tested) students

The second set of samples includes districts or schools that have unique reporting situations that require the implementation of a decision rule. This set is necessary to ensure that each rule is applied correctly.

The quality-assurance group uses a checklist to implement its procedures. Once the checklist is completed, sample reports are circulated for review by psychometric and program management staff. The appropriate sample reports are then sent to RIDE for review and signoff.

Chapter 5 PSYCHOMETRIC QUALITY

5.1 CLASSICAL ITEM ANALYSES

As noted in Brown (1983), “A test is only as good as the items it contains.” A complete evaluation of a test’s quality must include an evaluation of each item. Both *Standards for Educational and Psychological Testing* (AERA et al., 2014) and the *Code of Fair Testing Practices in Education* (Joint Committee on Testing Practices, 2004) include standards for identifying quality items. Items should predominantly assess the knowledge and skills that are identified as part of the domain being tested and should avoid assessing irrelevant factors. Items should also be unambiguous and free of grammatical errors, potentially insensitive content or language, and other confounding characteristics. In addition, items must not unfairly disadvantage students—in particular, racial, ethnic, or gender groups.

Both qualitative and quantitative analyses have been conducted to ensure that 2019 RICAS items meet these standards. For details on the qualitative analyses, please see the *2019 Next-Generation MCAS Technical Report*. This chapter presents statistical evaluations in four parts: (1) difficulty indices, (2) item-test correlations, (3) DIF statistics, and (4) dimensionality analyses. The item analyses presented here are based on the statewide administration of the RICAS assessments in spring 2019. Note that the information presented in this section is based only on the operational items, since those are the items on which student scores are calculated.

5.1.1 Classical Difficulty and Discrimination Indices

All selected-response and constructed-response items are evaluated in terms of item difficulty according to standard classical test theory practices. Difficulty is defined as the average proportion of points achieved on an item and is measured by obtaining the average score on an item and dividing it by the maximum possible score for the item. Selected-response items are scored dichotomously (correct vs. incorrect), so, for these items, the difficulty index is simply the proportion of students who correctly answered the item. Constructed-response items and essay items are scored polytomously, meaning that a student can achieve scores other than just 0 or 1 (e.g., 0, 1, 2, 3, or 4 for a 4-point constructed-response item). By computing the difficulty index as the average proportion of points achieved, the indices for the different item types are placed on a similar scale, ranging from 0.0 to 1.0 regardless of the item type. Although this index is traditionally described as a measure of difficulty, it is properly interpreted as an easiness index, because larger values indicate easier items. An index of 0.0 indicates that all students earned 0% of the item points, and an index of 1.0 indicates that all students received full credit for the item (i.e., all the item points). For additional details, please see the *2019 MCAS Next-Generation Technical Report*.

A summary of the item difficulty and item discrimination statistics for each grade and content area combination is presented in Table 5-1. Note that the statistics are presented for all items as well as separately by item type: selected response (SR), constructed response (CR), and essay (ES). The mean difficulty (p -value) and discrimination values shown in the table are within generally acceptable and expected ranges.

Table 5-1. Summary of Item Difficulty and Discrimination Statistics by Content Area and Grade

Content Area	Grade	Item Type	Number of Items	Difficulty		Discrimination		
				Mean	Standard Deviation	Mean	Standard Deviation	
ELA	3	ALL	26	0.58	0.16	0.43	0.13	
		SR	15	0.60	0.13	0.36	0.09	
		CR	7	0.63	0.17	0.45	0.08	
		ES	4	0.37	0.09	0.65	0.07	
	4	ALL	26	0.58	0.16	0.45	0.13	
		SR	15	0.63	0.13	0.37	0.09	
		CR	7	0.62	0.09	0.52	0.08	
		ES	4	0.31	0.04	0.64	0.02	
	5	ALL	28	0.60	0.16	0.44	0.13	
		SR	17	0.67	0.11	0.37	0.08	
		CR	5	0.65	0.15	0.42	0.06	
		ES	6	0.37	0.07	0.66	0.07	
	6	ALL	27	0.53	0.13	0.49	0.16	
		SR	15	0.58	0.08	0.40	0.08	
		CR	6	0.56	0.16	0.46	0.09	
		ES	6	0.37	0.10	0.75	0.02	
	7	ALL	27	0.53	0.14	0.46	0.17	
		SR	15	0.58	0.11	0.35	0.06	
		CR	6	0.57	0.12	0.47	0.09	
		ES	6	0.36	0.11	0.74	0.03	
	8	ALL	27	0.55	0.14	0.48	0.17	
		SR	15	0.62	0.10	0.39	0.06	
		CR	6	0.51	0.14	0.42	0.07	
		ES	6	0.42	0.12	0.77	0.02	
	Mathematics	3	ALL	40	0.59	0.19	0.48	0.09
			SR	17	0.70	0.14	0.42	0.07
			CR	23	0.51	0.17	0.52	0.08
		4	ALL	40	0.58	0.18	0.45	0.11
SR			18	0.65	0.18	0.42	0.12	
CR			22	0.52	0.17	0.48	0.11	
5		ALL	40	0.49	0.18	0.43	0.14	
		SR	18	0.53	0.18	0.34	0.12	
		CR	22	0.46	0.17	0.50	0.12	
6		ALL	40	0.48	0.20	0.44	0.12	
		SR	20	0.57	0.16	0.42	0.08	
		CR	20	0.39	0.20	0.47	0.15	
7		ALL	40	0.37	0.18	0.44	0.16	
		SR	16	0.48	0.16	0.32	0.15	
		CR	24	0.29	0.15	0.52	0.12	
8		ALL	40	0.46	0.16	0.45	0.14	
		SR	21	0.51	0.14	0.40	0.11	
		CR	19	0.40	0.15	0.50	0.16	

Caution should be exercised when comparing indices across grade levels. Differences may be due not only to differences in the item statistics on the test but may also be affected by differences in student abilities and/or differences in the standards and/or curricula taught in each grade.

Difficulty indices for selected-response items tend to be higher (indicating that students performed better on these items) than the difficulty indices for constructed-response items because selected-response items can be answered correctly by simply identifying rather than providing the correct answer, or by guessing. Similarly, discrimination indices for those constructed-response items with more than two points tend to be larger than those for dichotomous items because of the greater variability of the former (i.e., the partial credit these items allow). The restriction of range (i.e., only two score categories) in dichotomous items tends to make the discrimination indices lower. Note that these patterns are more consistent within item type, so when interpreting classical item statistics, comparisons should be emphasized among items of the same type.

In addition to the item difficulty and discrimination summaries presented above, item-level classical test theory statistics are provided in Appendix E. On RICAS items, the item difficulty and discrimination indices are within generally acceptable and expected ranges. Very few items were answered correctly at near-chance or near-perfect rates. Similarly, the positive discrimination indices indicate that students who performed well on individual items tended to perform well overall. There are a small number of items with discrimination indices below 0.20, but none were negative. While it is acceptable to include items with low discrimination values or with very high or very low item difficulty values when their content is needed to ensure that the content specifications are appropriately covered, there were very few such cases on the 2019 RICAS. Item-level score point distributions are provided for constructed-response items in Appendix F; for each item, the percentage of students who received each score point is presented.

5.1.2 Differential Item Functioning

For the RICAS spring 2019 administration, Differential Item Functioning (DIF) analyses were conducted for all subgroups (as defined in the No Child Left Behind Act) for which the sample size was adequate. Six subgroup comparisons were evaluated for DIF:

- male compared with female
- not ELL compared with ELL¹
- not economically disadvantaged compared with economically disadvantaged
- white compared with African American or Black
- white compared with Hispanic or Latino
- students with disabilities compared with students without disabilities

The tables in Appendix G present the number of items classified as either “low” or “high” DIF, in total and by group favored. The moderate number of items that exhibited low DIF and several that exhibited high DIF were reviewed by content and educational experts to rule out a source of bias prior to being included on the operational tests. For detailed information about how the DIF procedure was employed, please see the *2019 MCAS Next-Generation Technical Report*.

¹ ELL = English Language Learner (includes current and former English Language Learners).

5.1.3 Dimensionality Analysis

The purpose of dimensionality analysis is to investigate whether violation of the assumption of test unidimensionality is statistically detectable and, if so, (a) the degree to which unidimensionality is violated and (b) the nature of the multidimensionality.

The nonparametric IRT-based methods DIMTEST (Stout, 1987; Stout, Froelich, & Gao, 2001) and DETECT (Zhang & Stout, 1999) were applied to operational items for RICAS online test forms administered during the spring 2019 administrations. A total of 12 test forms were analyzed. The data for each grade were split into a training sample and a cross-validation sample. For all grades, there were over 10,200 student examinees per test form in both ELA and mathematics, so every training sample and cross-validation sample had at least 5,100 students. After randomly splitting the data into training and cross-validation samples, DIMTEST was applied to each data set to see if the null hypothesis of unidimensionality would be rejected. DETECT was then applied to each data set for which the DIMTEST null hypothesis was rejected in order to estimate the effect size of the multidimensionality. DETECT values less than 0.2 indicate very weak multidimensionality (or near unidimensionality); values of 0.2 to 0.4, weak to moderate multidimensionality; values of 0.4 to 1.0, moderate to strong multidimensionality; and values greater than 1.0, very strong multidimensionality (Roussos & Ozbek, 2006).

The results of the DIMTEST analyses indicated that the null hypothesis was rejected at a significance level of 0.01 for every data set. Because strict unidimensionality is an idealization that almost never holds exactly for a given data set, the statistical rejections in the DIMTEST results were not surprising. Indeed, because of the large sample sizes involved in the data sets, DIMTEST would be expected to be sensitive to even quite small violations of unidimensionality.

DETECT was then used to estimate the effect size for the violations of local independence for all the tests. Table 5-2 displays the multidimensionality effect-size estimates from DETECT.

Table 5-2. Multidimensionality Effect Sizes by Grade and Content Area

<i>Content Area</i>	<i>Grade</i>	<i>Multidimensionality Effect Size</i>
ELA	3	0.262
	4	0.220
	5	0.282
	6	0.306
	7	0.462
	8	0.441
	Average	0.329
Mathematics	3	0.196
	4	0.117
	5	0.152
	6	0.199
	7	0.153
	8	0.158
Average	0.163	

The DETECT values indicate weak or very weak multidimensionality for all the 2019 RICAS mathematics test forms. All the 2019 RICAS ELA test forms show weak to moderate multidimensionality.

The way in which DETECT divided the tests into clusters was also investigated to determine whether there were any discernable patterns with respect to the selected-response and constructed-response item types. Inspection of the DETECT clusters indicated that selected-response/constructed-response separation generally occurred much

more strongly with ELA than with mathematics. Specifically, for the ELA test forms, every grade had one set of clusters dominated by selected-response items and another set of clusters dominated by writing prompt items. On the mathematics test forms, there was less clear evidence of consistent separation of selected-response and constructed-response items.

In summary, for the 2019 dimensionality analyses, the violations of local independence, as evidenced by the DETECT effect sizes, were either very weak or weak in mathematics test forms, and were weak-to-moderate in ELA test forms. The patterns with respect to the selected-response and constructed-response items suggested that ELA tended to display more separation than mathematics.

5.2 RICAS RELIABILITY

5.2.1 Reliability and Standard Errors of Measurement

The approach that was implemented to assess the reliability of the 2019 RICAS tests was the α coefficient of Cronbach (1951). For details on the calculation of Cronbach's α coefficient, please see the *2019 MCAS Next-Generation Technical Report*. Table 5-3 presents descriptive statistics, Cronbach's α coefficient, and the raw score standard error of measurement (SEM) for each content area and grade. Statistics are based on operational items from online test forms, which were taken by most of the student examinee population. The reliability estimates range from 0.87 to 0.93, which are in generally acceptable ranges.

Table 5-3. Raw Score Descriptive Statistics, Cronbach's Alpha, and SEMs by Content Area and Grade

Content Area	Grade	Number of Students	Raw Score			Alpha (α)	SEM
			Maximum	Mean	Standard Deviation		
ELA	3	10,188	44	23.34	8.20	0.87	2.90
	4	10,410	44	22.80	8.45	0.89	2.80
	5	10,694	48	25.28	9.22	0.89	3.08
	6	10,792	51	23.62	10.69	0.91	3.21
	7	10,630	51	23.53	10.53	0.90	3.31
	8	10,491	51	24.74	10.76	0.91	3.28
Mathematics	3	10,321	48	24.10	11.69	0.93	3.00
	4	10,520	54	27.95	12.11	0.92	3.33
	5	10,807	54	23.79	12.05	0.91	3.59
	6	10,912	54	22.82	11.68	0.92	3.38
	7	10,735	54	18.96	11.24	0.91	3.37
	8	10,610	54	22.61	11.87	0.92	3.38

Because of the dependency of the α coefficients on the test-taking population and the test characteristics, precautions need to be taken when making inferences about the quality of one test by comparing its reliability to that of another test from a different grade or content area. To elaborate, reliability coefficients are highly influenced by test-taking population characteristics such as the range of individual differences in the group (i.e., variability within the population), average ability level of the population that took the exams, test designs, test difficulty, test length, ceiling or floor effect, and influence of guessing. Hence, "the reported reliability coefficient is only applicable to samples similar to that on which it was computed" (Anastasi & Urbina, 1997, p.107).

5.2.2 Reporting Subcategory Reliability

Reliabilities were calculated for the reporting subcategories within the 2019 RICAS content areas. Results and reporting category descriptions are presented in Appendix H. The reliability coefficients for the reporting subcategories range from 0.43 to 0.87, with a median of 0.72 and a standard deviation of 0.12. Lower reliabilities on subcategory scores are associated with very low numbers of items. Because they are based on a subset of items rather than the full test, subcategory reliabilities were typically lower than were overall test score reliabilities,

approximately to the degree expected based on the classical test theory (Haertel, 2006), and interpretations should take this into account. Qualitative differences among grades and content areas once again preclude valid inferences about the reliability of the full test score based on statistical comparisons among subtests.

5.2.3 Subgroup Reliability

The reliability coefficients discussed in the previous section were based on the overall population of students who took the 2019 RICAS online forms. Appendix H presents reliabilities for various subgroups of interest for ELA and mathematics, respectively. Cronbach's α coefficients were calculated based only on the members of the subgroup in question in the computations; values are calculated only for subgroups with 10 or more students. The reliability coefficients for subgroups range from 0.71 to 0.94 across the tests, with a median of 0.90 and a standard deviation of 0.032, indicating that reliabilities are generally within a reasonable range.

For several reasons, the subgroup reliability results should be interpreted with caution. Reliabilities are dependent not only on the measurement properties of a test but also on the statistical distribution of the studied subgroup. For example, subgroup sizes may vary considerably, which results in natural variation in reliability coefficients. Alternatively, α , which is a type of correlation coefficient, may be artificially depressed for subgroups with little variability (Draper & Smith, 1998). Moreover, there is no industry standard to interpret the strength of a reliability coefficient when the population of interest is a single subgroup.

5.2.4 Decision Accuracy and Consistency Results

Decision Accuracy and Consistency (DAC) analyses were conducted for online test forms at each performance achievement level. Results of the DAC analyses are provided in Tables 5-4 and 5-5 for the 2019 RICAS tests.

Table 5-4 includes overall accuracy indices with consistency indices displayed in parentheses next to the accuracy values, as well as overall kappa values. Overall ranges for accuracy (0.78–0.85), consistency (0.69–0.80), and kappa (0.53–0.66) indicate that most students were classified accurately and consistently with respect to measurement error and chance. Accuracy and consistency values conditional on achievement level are also given. For these calculations, the denominator is the proportion of students associated with a given achievement level. For example, the conditional accuracy value is 0.81 for *Not Meeting Expectations* for the grade 3 ELA test. This figure indicates that among the students whose true scores placed them in this classification, 81% would be expected to be in this classification when categorized according to their observed scores. Similarly, a consistency value of 0.67 indicates that 67% of students with observed scores in the *Not Meeting Expectations* level would be expected to score in this classification again if a second, parallel test form was taken.

For some testing situations, the greatest concern may be decisions around achievement level thresholds. In this case, accuracy at the *Partially Meeting Expectations/Meeting Expectations* threshold is critically important, which summarizes the percentage of students who are correctly classified either above or below the particular cutpoint. Table 5-4 provides the accuracy and consistency estimates and false positive and false negative decision rates at each cutpoint for the 2019 RICAS online tests. A false positive is the proportion of students whose observed scores were above the cut and whose true scores were below the cut. A false negative is the proportion of students whose observed scores were below the cut and whose true scores were above the cut.

In Table 5-5, the accuracy and consistency indices at the *Partially Meeting Expectations/Meeting Expectations* threshold range from 0.89–0.94 and 0.84–0.91, respectively. The false positive and false negative decision rates at the *Partially Meeting Expectations/Meeting Expectations* threshold both range from 3%–6%. These results indicate that nearly all students were correctly classified with respect to being above or below the *Partially Meeting Expectations/Meeting Expectations* cutpoint.

**Table 5-4. Summary of Decision Accuracy (and Consistency) Results
by Content Area and Grade—Overall and Conditional on Achievement Level**

Content Area	Grade	Overall	Kappa	Conditional on Achievement Level			
				Not Meeting Expectations	Partially Meeting Expectations	Meeting Expectations	Exceeding Expectations
ELA	3	0.78 (0.69)	0.53	0.81 (0.67)	0.79 (0.73)	0.77 (0.70)	0.72 (0.52)
	4	0.82 (0.74)	0.59	0.82 (0.68)	0.84 (0.80)	0.79 (0.71)	0.78 (0.59)
	5	0.82 (0.75)	0.59	0.81 (0.67)	0.84 (0.79)	0.81 (0.74)	0.78 (0.58)
	6	0.81 (0.73)	0.60	0.86 (0.78)	0.81 (0.75)	0.78 (0.71)	0.72 (0.55)
	7	0.81 (0.73)	0.60	0.85 (0.77)	0.81 (0.76)	0.78 (0.70)	0.70 (0.49)
	8	0.81 (0.73)	0.61	0.86 (0.78)	0.80 (0.74)	0.79 (0.72)	0.74 (0.56)
Mathematics	3	0.84 (0.78)	0.65	0.85 (0.76)	0.85 (0.81)	0.83 (0.78)	0.75 (0.58)
	4	0.85 (0.78)	0.65	0.84 (0.73)	0.85 (0.81)	0.84 (0.78)	0.79 (0.60)
	5	0.85 (0.79)	0.65	0.83 (0.70)	0.86 (0.81)	0.85 (0.79)	0.76 (0.52)
	6	0.85 (0.80)	0.66	0.84 (0.74)	0.86 (0.82)	0.86 (0.80)	0.77 (0.55)
	7	0.79 (0.72)	0.55	0.67 (0.55)	0.80 (0.74)	0.85 (0.79)	0.82 (0.67)
	8	0.84 (0.77)	0.63	0.80 (0.71)	0.85 (0.80)	0.84 (0.77)	0.79 (0.61)

Table 5-5. Summary of Decision Accuracy (and Consistency) Results by Content Area and Grade—Conditional on Cutpoint

Content Area	Grade	Not Meeting Expectations / Partially Meeting Expectations			Partially Meeting Expectations / Meeting Expectations			Meeting Expectations / Exceeding Expectations		
		Accuracy (consistency)	False		Accuracy (consistency)	False		Accuracy (consistency)	False	
			Positive	Negative		Positive	Negative		Positive	Negative
ELA	3	0.95 (0.93)	0.02	0.03	0.89 (0.84)	0.06	0.06	0.94 (0.92)	0.04	0.02
	4	0.95 (0.92)	0.02	0.04	0.90 (0.86)	0.05	0.04	0.97 (0.96)	0.02	0.01
	5	0.95 (0.93)	0.02	0.03	0.90 (0.86)	0.05	0.05	0.97 (0.96)	0.02	0.01
	6	0.94 (0.91)	0.03	0.04	0.91 (0.88)	0.04	0.04	0.95 (0.94)	0.03	0.02
	7	0.93 (0.90)	0.03	0.04	0.91 (0.88)	0.05	0.04	0.97 (0.96)	0.02	0.01
	8	0.93 (0.90)	0.03	0.04	0.92 (0.88)	0.04	0.04	0.96 (0.95)	0.03	0.01
Mathematics	3	0.96 (0.94)	0.02	0.03	0.92 (0.89)	0.04	0.04	0.96 (0.95)	0.02	0.01
	4	0.95 (0.93)	0.02	0.03	0.92 (0.88)	0.04	0.04	0.98 (0.97)	0.02	0.01
	5	0.95 (0.93)	0.02	0.03	0.91 (0.88)	0.05	0.04	0.99 (0.98)	0.01	0.00
	6	0.95 (0.92)	0.02	0.03	0.92 (0.89)	0.04	0.04	0.99 (0.98)	0.01	0.00
	7	0.88 (0.84)	0.06	0.07	0.93 (0.90)	0.04	0.03	0.99 (0.98)	0.01	0.00
	8	0.92 (0.88)	0.04	0.05	0.94 (0.91)	0.04	0.03	0.98 (0.98)	0.01	0.00

The indices above are derived from Livingston and Lewis's (1995) method of estimating DAC. Livingston and Lewis discuss two versions of the accuracy and consistency tables. A standard version performs calculations for forms parallel to the form taken. An "adjusted" version adjusts the results of one form to match the observed score distribution obtained in the data. The tables use the standard version for two reasons: (1) This "unadjusted" version can be considered a smoothing of the data, thereby decreasing the variability of the results; and (2) for results dealing with the consistency of two parallel forms, the unadjusted tables are symmetrical, indicating that the two parallel forms have the same statistical properties. This second reason is consistent with the notion of forms that are parallel (i.e., it is more intuitive and interpretable for two parallel forms to have the same statistical distribution).

As with other methods of evaluating reliability, DAC statistics that are calculated based on small groups can be expected to be lower than those calculated based on larger groups. For this reason, the values presented in Tables 5-4 and 5-5 should be interpreted with caution. In addition, it is important to remember that it might be inappropriate to compare DAC statistics across grades and content areas.

REFERENCES

- Allen, M. J., & Yen, W. M. (1979). *Introduction to measurement theory*. Belmont, CA: Wadsworth, Inc.
- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2014). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.
- Anastasi, A., & Urbina, S. (1997). *Psychological testing* (7th ed.). Upper Saddle River, NJ: Prentice-Hall.
- Brown, F. G. (1983). *Principles of educational and psychological testing* (3rd ed.). Fort Worth, TX: Holt, Rinehart and Winston.
- Charter, R. A. (1999). Sample size requirements for precise estimates of reliability, generalizability, and validity coefficients. *Journal of Clinical and Experimental Neuropsychology*, 21(4), 559–566(8).
- Chicago Manual of Style* (16th ed.). (2003). Chicago: University of Chicago Press.
- Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, 20, 37–46
- Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika* 16, 297–334.
- Draper, N. R., & Smith, H. (1998). *Applied regression analysis* (3rd ed.). New York, NY: John Wiley and Sons, Inc.
- Haertel, E. H. (2006). Reliability. In R.L. Brennan (Ed). *Educational measurement* (pp. 65-110). Westport, CT: Praeger Publishers.
- Joint Committee on Testing Practices. (2004). *Code of fair testing practices in education*. Washington, DC: Joint Committee on Testing Practices. Available from apa.org/science/programs/testing/fair-code.aspx.
- Livingston, S. A., & Lewis, C. (1995). Estimating the consistency and accuracy of classifications based on test scores. *Journal of Educational Measurement* 32, 179–197.
- Roussos, L. A., & Ozbek, O. Y. (2006). Formulation of the DETECT population parameter and evaluation of DETECT estimator bias. *Journal of Educational Measurement* 43, 215–243.
- Stout, W. F. (1987). A nonparametric approach for assessing latent trait dimensionality. *Psychometrika* 52, 589–617.
- Stout, W. F., Froelich, A. G., & Gao, F. (2001). Using resampling methods to produce an improved DIMTEST procedure. In A. Boomsma, M. A. J. van Duijn, & T. A. B. Snijders (Eds.), *Essays on Item Response Theory* (pp. 357–375). New York, NY: Springer-Verlag.
- Zhang, J., & Stout, W. F. (1999). The theoretical DETECT index of dimensionality and its application to approximate simple structure. *Psychometrika* 64, 213–249.

APPENDICES

APPENDIX A
PARTICIPATION RATES

**Table A-1. Summary of Participation by Student Subgroup
English Language Arts, Grades 3–8**

<i>Description</i>	<i>Number Tested</i>	<i>Percent Tested</i>
All Students	63,205	100
ELL	9,650	15.27
Economically Disadvantaged	31,089	49.19
African American	5,442	8.61
Asian	2,065	3.27
Hispanic	16,540	26.17
Native American/Alaska Native	458	0.72
White	35,617	56.35
Pacific Islander/Hawaiian	95	0.15
Multiracial	2,942	4.65
Male	32,300	51.10
Female	30,859	48.82
Special Education	9,329	14.76

**Table A-2. Summary of Participation by Student Subgroup
Mathematics, Grades 3–8**

<i>Description</i>	<i>Number Tested</i>	<i>Percent Tested</i>
All Students	63,905	100
ELL	10,347	16.19
Economically Disadvantaged	31,569	49.40
African American	5,521	8.64
Asian	2,107	3.30
Hispanic	17,032	26.65
Native American/Alaska Native	460	0.72
White	35,696	55.86
Pacific Islander/Hawaiian	96	0.15
Multiracial	2,947	4.61
Male	32,677	51.13
Female	31,182	48.79
Special Education	9,346	14.62

APPENDIX B

ACCOMMODATIONS

Table B-1. Numbers of Students Tested with and Without Accommodations by Content Area and Grade

Content Area	Grade	Number of Students Tested	
		With Accommodations	Without Accommodations
ELA	3	935	9,253
	4	1,085	9,325
	5	1,077	9,617
	6	944	9,848
	7	850	9,780
	8	779	9,712
Mathematics	3	2,184	8,137
	4	2,189	8,331
	5	2,065	8,742
	6	1,526	9,386
	7	1,380	9,355
	8	1,227	9,383

Table B-2. Numbers of Students Tested with Accommodations by Accommodation Type and Grade—ELA

Description	Grade 3	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8
Color Contrast	10	17	15	11	6	5
Black on Cream	8	6	13	5	2	3
Black on Light Blue	1	5	2	3	1	0
Black on Light Magenta	1	1	0	1	0	0
White on Black	0	4	0	2	3	0
Yellow on Blue	0	1	0	0	0	0
Dark Gray on Pale Green	0	0	0	0	0	2
Answer Masking	81	161	147	56	15	37
Large Print Test Edition	1	0	1	0	1	3
Screen Reader Edition	0	1	0	0	0	0
Assistive Technology	8	13	11	3	6	4
Braille Test Edition	0	0	0	0	0	0
Human Read Aloud as a Non-Standard Accommodation	34	34	24	7	18	16
Human Signer as a Standard Accommodation	2	6	6	3	4	2
Human Signer as a Non-Standard Accommodation	0	0	1	0	0	0
Text-to-Speech	114	116	97	135	98	126
Human Scribe as a Non-Standard Accommodation	37	41	32	21	13	10
Speech-to-Text as a Non-Standard Accommodation	40	37	45	23	24	8
Typed Responses	0	0	0	0	0	1
Spell-checker	30	33	64	27	38	30
Word Prediction	20	24	29	25	16	14
Graphic Organizer/Reference Sheet	607	687	676	550	423	397
Any Other accommodations	175	243	195	199	185	157
Bilingual Dictionary and Glossary	90	83	120	207	237	199

**Table B-3. Numbers of Students Tested with Accommodations
by Accommodation Type and Grade—Mathematics**

<i>Description</i>	<i>Grade 3</i>	<i>Grade 4</i>	<i>Grade 5</i>	<i>Grade 6</i>	<i>Grade 7</i>	<i>Grade 8</i>
Color Contrast	10	13	14	9	8	5
Black on Cream	8	4	12	3	3	2
Black on Light Blue	1	4	2	2	1	0
Black on Light Magenta	1	1	0	1	0	1
White on Black	0	3	0	3	4	0
Yellow on Blue	0	1	0	0	0	0
Dark Gray on Pale Green	0	0	0	0	0	2
Answer Masking	82	153	140	49	14	36
Large Print Test Edition	1	0	1	0	2	3
Screen Reader Edition	0	0	0	0	0	0
Assistive Technology	1	6	7	7	18	8
Braille Test Edition	0	0	0	0	0	0
Human Read Aloud as a Standard Accommodation	121	103	68	11	18	20
Human Signer as a Standard Accommodation	2	6	6	3	4	2
Text-to-Speech	1,814	1,700	1,577	1,037	757	655
Human Scribe as a Standard Accommodation	27	49	33	18	11	7
Speech-to-Text as a Standard	28	27	33	16	18	8
Typed Responses	0	0	0	0	0	1
Calculation Device on Non-Calculator Session	82	88	142	225	256	244
Graphic Organizer/Reference Sheet	546	657	614	526	432	425
Any Other accommodations	182	228	185	199	185	164
Spanish	105	123	116	150	158	157
Bilingual Dictionary and Glossary	108	102	142	240	216	205

APPENDIX C
INTERRATER CONSISTENCY

Table C-1. Item-Level Interrater Consistency Statistics—ELA Grade 3

Item Number	Number of		Percent		Correlation	Percent of Third Scores	Kappa
	Score Categories	Responses Scored Twice	Exact	Adjacent			
EL293264#SCORE_TRAIT_Conv	4	970	63.81	34.74	0.79	2.89	0.655
EL293264#SCORE_TRAIT_Ideadev	5	970	63.20	35.26	0.82	2.89	0.669
EL708642952	4	986	73.63	25.86	0.83	0.51	0.732
EL715954244#SCORE_TRAIT_Conv	4	983	80.16	19.63	0.71	0.61	0.626
EL715954244#SCORE_TRAIT_Ideadev	5	983	82.40	17.09	0.76	0.61	0.692

Table C-2. Item-Level Interrater Consistency Statistics—ELA Grade 4

Item Number	Number of		Percent		Correlation	Percent of Third Scores	Kappa
	Score Categories	Responses Scored Twice	Exact	Adjacent			
EL707254137#SCORE_TRAIT_Conv	4	995	74.77	24.32	0.76	1.31	0.664
EL707254137#SCORE_TRAIT_Ideadev	5	995	75.38	24.22	0.73	1.31	0.626
EL710438990#SCORE_TRAIT_Conv	4	1,019	73.31	26.30	0.83	1.28	0.716
EL710438990#SCORE_TRAIT_Ideadev	5	1,019	63.30	35.82	0.80	1.28	0.644
EL712438196	4	1,028	79.86	19.75	0.84	0.39	0.758

Table C-3. Item-Level Interrater Consistency Statistics—ELA Grade 5

Item Number	Number of		Percent		Correlation	Percent of Third Scores	Kappa
	Score Categories	Responses Scored Twice	Exact	Adjacent			
EL624182427#SCORE_TRAIT_Conv	4	1,028	72.37	25.19	0.76	3.50	0.667
EL624182427#SCORE_TRAIT_Ideadev	5	1,028	75.39	22.86	0.82	3.50	0.732
EL709062207#SCORE_TRAIT_Conv	4	1,046	71.70	28.01	0.82	1.05	0.696
EL709062207#SCORE_TRAIT_Ideadev	5	1,046	73.42	25.53	0.82	1.05	0.689
EL709229186#SCORE_TRAIT_Conv	4	1,060	64.53	34.43	0.82	4.43	0.632
EL709229186#SCORE_TRAIT_Ideadev	5	1,060	52.55	43.49	0.83	4.43	0.609

Table C-4. Item-Level Interrater Consistency Statistics—ELA Grade 6

Item Number	Number of		Percent		Correlation	Percent of Third Scores	Kappa
	Score Categories	Responses Scored Twice	Exact	Adjacent			
EL707351199#SCORE_TRAIT_Conv	4	1,030	73.50	25.92	0.86	1.46	0.751
EL707351199#SCORE_TRAIT_Ideadev	6	1,030	64.27	34.76	0.85	1.46	0.706
EL710355409#SCORE_TRAIT_Conv	4	1,043	65.68	33.56	0.79	2.11	0.658
EL710355409#SCORE_TRAIT_Ideadev	6	1,043	73.06	25.41	0.84	2.11	0.729
EL712756190#SCORE_TRAIT_Conv	4	1,044	73.56	25.48	0.85	1.82	0.740
EL712756190#SCORE_TRAIT_Ideadev	6	1,044	69.73	28.93	0.85	1.82	0.711

Table C-5. Item-Level Interrater Consistency Statistics—ELA Grade 7

Item Number	Number of		Percent		Correlation	Percent of Third Scores	Kappa
	Score Categories	Responses Scored Twice	Exact	Adjacent			
EL285797#SCORE_TRAIT_Conv	4	1,024	74.02	25.98	0.83	0.00	0.720
EL285797#SCORE_TRAIT_Ideadev	6	1,024	75.39	24.61	0.82	0.00	0.711
EL707935717#SCORE_TRAIT_Conv	4	1,045	73.49	26.32	0.87	0.96	0.759
EL707935717#SCORE_TRAIT_Ideadev	6	1,045	70.05	29.19	0.88	0.96	0.739
EL714343909#SCORE_TRAIT_Conv	4	996	75.80	23.90	0.89	0.70	0.796
EL714343909#SCORE_TRAIT_Ideadev	6	996	71.39	28.11	0.92	0.70	0.794

Table C-6. Item-Level Interrater Consistency Statistics—ELA Grade 8

Item Number	Number of		Percent		Correlation	Percent of Third Scores	Kappa
	Score Categories	Responses Scored Twice	Exact	Adjacent			
EL709130565#SCORE_TRAIT_Conv	4	1,015	77.44	22.17	0.90	1.77	0.812
EL709130565#SCORE_TRAIT_Ideadev	6	1,015	68.47	29.95	0.90	1.77	0.770
EL709184717#SCORE_TRAIT_Conv	4	1,017	76.99	22.62	0.89	1.18	0.794
EL709184717#SCORE_TRAIT_Ideadev	6	1,017	74.43	24.78	0.89	1.18	0.784
EL714447652#SCORE_TRAIT_Conv	4	983	78.03	21.87	0.88	2.24	0.791
EL714447652#SCORE_TRAIT_Ideadev	6	983	80.06	17.80	0.90	2.24	0.827

Table C-7. Item-Level Interrater Consistency Statistics—Mathematics Grade 3

Item Number	Number of		Percent		Correlation	Percent of Third Scores	Kappa
	Score Categories	Responses Scored Twice	Exact	Adjacent			
MA283013A	4	1,011	93.37	5.93	0.96	0.69	0.929
MA283013A_ES	4	10	100.00	0.00	1.00	0.00	1.000
MA703084007	4	1,115	94.53	5.47	0.97	0.00	0.938
MA703084007_ES	4	10	90.00	10.00		0.00	
MA703131543	4	957	88.19	10.87	0.94	0.94	0.883
MA703131543_ES	4	11	100.00	0.00		0.00	
MA724333304	4	998	96.39	3.51	0.99	0.10	0.972
MA724333304_ES	4	10	90.00	10.00	0.96	0.00	0.904

Table C-8. Item-Level Interrater Consistency Statistics—Mathematics Grade 4

Item Number	Number of		Percent		Correlation	Percent of Third Scores	Kappa
	Score Categories	Responses Scored Twice	Exact	Adjacent			
MA293812	5	935	89.41	9.63	0.94	0.96	0.901
MA293812_ES	5	11	100.00	0.00	1.00	0.00	1.000
MA307317	5	933	83.49	14.90	0.92	1.61	0.856
MA307317_ES	5	10	90.00	10.00	0.93	0.00	0.853
MA311583	5	980	91.84	7.96	0.96	0.20	0.925
MA311583_ES	5	12	91.67	8.33	0.95	0.00	0.902
MA713939739	5	1,073	85.18	13.23	0.95	1.58	0.898
MA713939739_ES	5	11	90.91	9.09	0.96	0.00	0.909

Table C-9. Item-Level Interrater Consistency Statistics—Mathematics Grade 5

Item Number	Number of		Percent		Correlation	Percent of Third Scores	Kappa
	Score Categories	Responses Scored Twice	Exact	Adjacent			
MA280511	5	965	79.48	19.27	0.91	1.24	0.834
MA280511_ES	5	11	100.00	0.00	1.00	0.00	1.000
MA303765	5	1,060	89.53	9.25	0.97	1.23	0.931
MA303765_ES	5	10	100.00	0.00	1.00	0.00	1.000
MA306457	5	1,128	92.20	7.71	0.98	0.09	0.956
MA306457_ES	5	12	100.00	0.00	1.00	0.00	1.000
MA715102462	5	1,007	86.00	12.71	0.94	1.29	0.886
MA715102462_ES	5	9	77.78	22.22	0.82	0.00	0.571

Table C-10. Item-Level Interrater Consistency Statistics—Mathematics Grade 6

Item Number	Number of		Percent		Correlation	Percent of Third Scores	Kappa
	Score Categories	Responses Scored Twice	Exact	Adjacent			
MA290253	5	1,102	90.93	8.98	0.97	0.09	0.927
MA290253_ES	5	10	100.00	0.00	1.00	0.00	1.000
MA703251109	5	1,040	88.37	11.06	0.96	0.48	0.917
MA703251109_ES	5	12	91.67	8.33	0.77	0.00	0.750
MA703253363	5	1,121	86.08	12.93	0.95	0.98	0.890
MA703253363_ES	5	13	92.31	7.69	0.96	0.00	0.893
MA713935781	5	1,184	93.58	6.08	0.97	0.34	0.939
MA713935781_ES	5	13	100.00	0.00		0.00	

Table C-11. Item-Level Interrater Consistency Statistics—Mathematics Grade 7

Item Number	Number of		Percent		Correlation	Percent of Third Scores	Kappa
	Score Categories	Responses Scored Twice	Exact	Adjacent			
MA302339	5	1,077	87.09	12.53	0.96	0.37	0.913
MA302339_ES	5	15	100.00	0.00	1.00	0.00	1.000
MA306566	5	1,107	87.62	11.74	0.95	0.63	0.890
MA306566_ES	5	14	100.00	0.00	1.00	0.00	1.000
MA703943771	5	963	93.35	5.71	0.97	0.93	0.942
MA703943771_ES	5	14	100.00	0.00		0.00	
MA715009326	5	894	91.61	7.72	0.96	0.67	0.932
MA715009326_ES	5	14	100.00	0.00	1.00	0.00	1.000

Table C-12. Item-Level Interrater Consistency Statistics—Mathematics Grade 8

Item Number	Number of		Percent		Correlation	Percent of Third Scores	Kappa
	Score Categories	Responses Scored Twice	Exact	Adjacent			
MA301680	5	866	93.88	6.12	0.97	0.00	0.946
MA301680_ES	5	13	100.00	0.00	1.00	0.00	1.000
MA311433	5	1,054	83.02	16.98	0.92	0.00	0.851
MA311433_ES	5	16	100.00	0.00	1.00	0.00	1.000
MA704872840	5	1,049	89.80	10.20	0.98	0.00	0.942
MA704872840_ES	5	11	100.00	0.00	1.00	0.00	1.000

APPENDIX D

ACHIEVEMENT LEVEL DISTRIBUTIONS

**Table D-1. Cut Scores on the Theta Metric and Reporting Scale
by Content Area and Grade**

Content Area	Grade	Theta			Scaled Score				
		Cut 1	Cut 2	Cut 3	Min	Cut 1	Cut 2	Cut 3	Max
ELA	3	-1.581	0.011	1.604	440	470	500	530	560
	4	-1.561	0.031	1.623	440	470	500	530	560
	5	-1.659	0.038	1.734	440	470	500	530	560
	6	-1.591	-0.011	1.570	440	470	500	530	560
	7	-1.560	0.011	1.582	440	470	500	530	560
	8	-1.456	0.051	1.559	440	470	500	530	560
Mathematics	3	-1.377	0.027	1.432	440	470	500	530	560
	4	-1.379	0.054	1.487	440	470	500	530	560
	5	-1.551	0.025	1.601	440	470	500	530	560
	6	-1.518	-0.008	1.502	440	470	500	530	560
	7	-1.414	0.031	1.476	440	470	500	530	560
	8	-1.496	-0.008	1.479	440	470	500	530	560

Table D-2. Achievement-Level Distributions by Grade—ELA

Grade	Achievement Level	Percent in Level	
		2019	2018
3	Not Meeting Expectations	11.55	11.12
	Partially Meeting Expectations	40.55	48.62
	Meeting Expectations	40.07	34.89
	Exceeding Expectations	7.83	5.37
4	Not Meeting Expectations	14.24	14.82
	Partially Meeting Expectations	48.52	47.04
	Meeting Expectations	33.60	34.07
	Exceeding Expectations	3.64	4.07
5	Not Meeting Expectations	12.58	13.70
	Partially Meeting Expectations	48.33	48.85
	Meeting Expectations	35.35	35.16
	Exceeding Expectations	3.75	2.29
6	Not Meeting Expectations	20.95	19.80
	Partially Meeting Expectations	39.94	45.82
	Meeting Expectations	32.64	30.54
	Exceeding Expectations	6.48	3.84
7	Not Meeting Expectations	22.94	28.66
	Partially Meeting Expectations	45.47	47.44
	Meeting Expectations	27.36	21.67
	Exceeding Expectations	4.23	2.23
8	Not Meeting Expectations	23.21	30.30
	Partially Meeting Expectations	40.48	41.53
	Meeting Expectations	30.48	25.11
	Exceeding Expectations	5.82	3.06

Table D-3. Achievement-Level Distributions by Grade—Mathematics

<i>Grade</i>	<i>Achievement Level</i>	<i>Percent in Level</i>	
		<i>2019</i>	<i>2018</i>
3	Not Meeting Expectations	20.49	19.72
	Partially Meeting Expectations	43.40	44.89
	Meeting Expectations	31.12	31.15
	Exceeding Expectations	4.99	4.24
4	Not Meeting Expectations	20.34	23.48
	Partially Meeting Expectations	47.10	49.69
	Meeting Expectations	29.18	24.62
	Exceeding Expectations	3.37	2.21
5	Not Meeting Expectations	18.68	20.76
	Partially Meeting Expectations	51.19	52.45
	Meeting Expectations	28.22	25.46
	Exceeding Expectations	1.91	1.34
6	Not Meeting Expectations	19.03	20.56
	Partially Meeting Expectations	53.02	54.19
	Meeting Expectations	25.31	24.09
	Exceeding Expectations	2.63	1.17
7	Not Meeting Expectations	22.65	25.65
	Partially Meeting Expectations	49.69	47.38
	Meeting Expectations	24.71	25.16
	Exceeding Expectations	2.95	1.81
8	Not Meeting Expectations	24.25	23.86
	Partially Meeting Expectations	51.28	53.22
	Meeting Expectations	21.91	21.50
	Exceeding Expectations	2.55	1.42

APPENDIX E

ITEM-LEVEL CLASSICAL STATISTICS

Table E-1. Item-Level Classical Test Theory Statistics—ELA Grade 3

<i>Item</i>		<i>Difficulty</i>	<i>Discrimination</i>	<i>Percent Omitted (%)</i>
<i>Number</i>	<i>Type</i>			
EL293236	SR	0.77	0.35	0
EL293239	SR	0.76	0.44	0
EL293246	CR	0.70	0.53	0
EL293248	SR	0.81	0.44	0
EL293252	SR	0.65	0.31	0
EL293257	SR	0.76	0.36	0
EL293259	CR	0.92	0.49	1
EL293264#SCORE_TRAIT_Conv	ES	0.48	0.71	1
EL293264#SCORE_TRAIT_Ideadev	ES	0.37	0.70	1
EL6245000741	SR	0.54	0.19	0
EL624585614	CR	0.68	0.45	0
EL62458579	CR	0.65	0.56	0
EL6246038101	SR	0.65	0.40	0
EL624605530	SR	0.57	0.23	0
EL624606184	SR	0.58	0.43	0
EL715954244#SCORE_TRAIT_Conv	ES	0.37	0.58	1
EL715954244#SCORE_TRAIT_Ideadev	ES	0.27	0.59	1
EL716049700	SR	0.45	0.42	0
EL708637981	SR	0.44	0.22	0
EL708638799	CR	0.55	0.38	0
EL708639108	CR	0.56	0.35	0
EL708639794	SR	0.62	0.41	0
EL708640847	SR	0.41	0.39	0
EL708641332	SR	0.55	0.46	0
EL708642262	SR	0.51	0.35	0
EL708642952	CR	0.38	0.42	1

Table E-2. Item-Level Classical Test Theory Statistics—ELA Grade 4

<i>Item</i>		<i>Difficulty</i>	<i>Discrimination</i>	<i>Percent Omitted (%)</i>
<i>Number</i>	<i>Type</i>			
EL312475	SR	0.58	0.33	0
EL312477	SR	0.77	0.45	0
EL312493	SR	0.86	0.36	0
EL312500	SR	0.74	0.39	0
EL312506	SR	0.48	0.32	0
EL707245731	SR	0.45	0.20	0
EL707248916	SR	0.59	0.33	0
EL707249140	SR	0.66	0.45	0
EL707249352	CR	0.73	0.51	0
EL707249695	SR	0.65	0.43	0
EL707252385	SR	0.80	0.48	0
EL707253226	CR	0.67	0.52	0
EL707254137#SCORE_TRAIT_Conv	ES	0.35	0.64	0
EL707254137#SCORE_TRAIT_Ideadev	ES	0.29	0.64	0
EL710438990#SCORE_TRAIT_Conv	ES	0.34	0.65	0
EL710438990#SCORE_TRAIT_Ideadev	ES	0.27	0.61	0
EL710453346	CR	0.67	0.34	0
EL710740005	CR	0.49	0.52	0
EL712435018	CR	0.69	0.57	0
EL712438196	CR	0.50	0.57	0
EL712440994	SR	0.69	0.50	0
EL712444620	CR	0.62	0.59	0
EL712446921	SR	0.60	0.39	0
EL712446949	SR	0.45	0.18	0
EL712447272	SR	0.63	0.38	0
EL712465873	SR	0.43	0.43	0

Table E-3. Item-Level Classical Test Theory Statistics—ELA Grade 5

<i>Item</i>		<i>Difficulty</i>	<i>Discrimination</i>	<i>Percent Omitted (%)</i>
<i>Number</i>	<i>Type</i>			
EL709062113	CR	0.63	0.41	0
EL709062171	CR	0.59	0.44	0
EL709062207#SCORE_TRAIT_Conv	ES	0.38	0.67	0
EL709062207#SCORE_TRAIT_Ideadev	ES	0.30	0.70	0
EL709062416	SR	0.65	0.28	0
EL709062438	SR	0.64	0.35	0
EL709062866	SR	0.52	0.47	*
EL709068950	SR	0.78	0.50	*
EL709068976	SR	0.74	0.45	0
EL709069021	SR	0.48	0.34	0
EL709229186#SCORE_TRAIT_Conv	ES	0.47	0.64	0
EL709229186#SCORE_TRAIT_Ideadev	ES	0.31	0.53	0
EL709236062	SR	0.80	0.41	0
EL709236177	SR	0.67	0.43	*
EL709236521	SR	0.61	0.30	0
EL709237137	CR	0.80	0.34	0
EL709237220	SR	0.68	0.35	0
EL709237264	SR	0.83	0.35	0
EL712953695	SR	0.61	0.37	0
EL624175088	SR	0.61	0.40	0
EL624178677	SR	0.69	0.37	0
EL624180157	SR	0.51	0.21	0
EL624182427#SCORE_TRAIT_Conv	ES	0.45	0.70	0
EL624182427#SCORE_TRAIT_Ideadev	ES	0.32	0.69	0
EL627148548	SR	0.66	0.47	*
EL733854336	CR	0.44	0.39	0
EL733939052	SR	0.84	0.32	*
EL733940008	CR	0.79	0.50	0

* blank values represent no omitted responses on an item, and 0% is a result of rounding for very small values.

Table E-4. Item-Level Classical Test Theory Statistics—ELA Grade 6

<i>Item</i>		<i>Difficulty</i>	<i>Discrimination</i>	<i>Percent Omitted (%)</i>
<i>Number</i>	<i>Type</i>			
EL707540506	SR	0.58	0.40	*
EL707540543	SR	0.57	0.50	0
EL707541984	CR	0.28	0.29	0
EL707542115	CR	0.52	0.40	0
EL707543302	SR	0.71	0.36	0
EL710355409#SCORE_TRAIT_Conv	ES	0.50	0.76	1
EL710355409#SCORE_TRAIT_Ideadev	ES	0.32	0.73	1
EL710366412	SR	0.50	0.36	0
EL711077555	SR	0.54	0.32	0
EL712756190#SCORE_TRAIT_Conv	ES	0.41	0.76	1
EL712756190#SCORE_TRAIT_Ideadev	ES	0.26	0.74	1
EL712828407	SR	0.65	0.49	0
EL712828536	SR	0.65	0.34	0
EL712828842	SR	0.47	0.29	0
EL712829123	SR	0.67	0.43	0
EL712829207	SR	0.62	0.49	0
EL720341012	CR	0.70	0.53	0
EL720544333	CR	0.55	0.48	0
EL707347158	SR	0.64	0.47	*
EL707351199#SCORE_TRAIT_Conv	ES	0.46	0.77	1
EL707351199#SCORE_TRAIT_Ideadev	ES	0.28	0.73	1
EL707642910	SR	0.46	0.24	0
EL707643593	CR	0.58	0.51	*
EL710139135	SR	0.62	0.46	*
EL710269684	CR	0.71	0.52	*
EL710731348	SR	0.49	0.35	0
EL710731869	SR	0.54	0.46	0

* blank values represent no omitted responses on an item, and 0% is a result of rounding for very small values.

Table E-5. Item-Level Classical Test Theory Statistics—ELA Grade 7

<i>Item</i>		<i>Difficulty</i>	<i>Discrimination</i>	<i>Percent Omitted (%)</i>
<i>Number</i>	<i>Type</i>			
EL285787	SR	0.64	0.39	0
EL285795	SR	0.67	0.44	0
EL285797#SCORE_TRAIT_Conv	ES	0.41	0.77	1
EL285797#SCORE_TRAIT_Ideadev	ES	0.26	0.75	1
EL707803060	SR	0.58	0.33	*
EL707886390	SR	0.54	0.35	0
EL707931187	CR	0.52	0.37	0
EL707931508	SR	0.73	0.38	0
EL707931806	SR	0.51	0.39	0
EL707934528	SR	0.45	0.25	0
EL707935717#SCORE_TRAIT_Conv	ES	0.52	0.73	1
EL707935717#SCORE_TRAIT_Ideadev	ES	0.34	0.72	1
EL711859903	CR	0.72	0.36	*
EL718059830	CR	0.56	0.50	0
EL718150308	CR	0.46	0.52	0
EL722257627	SR	0.70	0.43	0
EL723051363	SR	0.63	0.39	0
EL729724758	SR	0.46	0.24	0
EL714338029	SR	0.46	0.30	0
EL71433887	SR	0.73	0.38	0
EL714341461	SR	0.51	0.29	0
EL714341617	SR	0.66	0.32	0
EL714343321	SR	0.44	0.39	0
EL714343909#SCORE_TRAIT_Conv	ES	0.39	0.77	1
EL714343909#SCORE_TRAIT_Ideadev	ES	0.23	0.70	1
EL719947920	CR	0.45	0.49	0
EL723042397	CR	0.73	0.58	0

* blank values represent no omitted responses on an item, and 0% is a result of rounding for very small values.

Table E-6. Item-Level Classical Test Theory Statistics—ELA Grade 8

<i>Item</i>		<i>Difficulty</i>	<i>Discrimination</i>	<i>Percent Omitted (%)</i>
<i>Number</i>	<i>Type</i>			
EL708953188	SR	0.65	0.34	0
EL708956273	CR	0.41	0.33	0
EL708963785	SR	0.81	0.37	0
EL708970664	SR	0.47	0.43	0
EL709036142	SR	0.68	0.33	0
EL709042787	SR	0.70	0.36	0
EL709072879	CR	0.33	0.37	0
EL709130565#SCORE_TRAIT_Conv	ES	0.57	0.78	1
EL709130565#SCORE_TRAIT_Ideadev	ES	0.37	0.78	1
EL709162988	SR	0.49	0.32	0
EL709170532	SR	0.52	0.33	0
EL709171912	SR	0.67	0.40	0
EL709172748	SR	0.54	0.52	*
EL709173329	SR	0.70	0.43	0
EL709184717#SCORE_TRAIT_Conv	ES	0.53	0.76	1
EL709184717#SCORE_TRAIT_Ideadev	ES	0.37	0.74	1
EL713050542	CR	0.56	0.45	0
EL728531887	CR	0.64	0.39	0
EL714442355	SR	0.65	0.31	0
EL714443470	CR	0.43	0.48	0
EL714444644	SR	0.65	0.42	0
EL714444869	SR	0.52	0.42	0
EL714445444	SR	0.68	0.40	*
EL714447304	CR	0.68	0.51	0
EL714447652#SCORE_TRAIT_Conv	ES	0.43	0.79	2
EL714447652#SCORE_TRAIT_Ideadev	ES	0.24	0.77	2
EL715851641	SR	0.62	0.42	0

* blank values represent no omitted responses on an item, and 0% is a result of rounding for very small values.

Table E-7. Item-Level Classical Test Theory Statistics—Mathematics Grade 3

<i>Item</i>		<i>Difficulty</i>	<i>Discrimination</i>	<i>Percent Omitted (%)</i>
<i>Number</i>	<i>Type</i>			
MA207011	SR	0.72	0.48	*
MA300745	CR	0.54	0.54	0
MA310866	SR	0.88	0.36	*
MA260963	SR	0.85	0.43	0
MA260965	SR	0.31	0.26	*
MA283013A	CR	0.33	0.54	0
MA300747	CR	0.64	0.53	0
MA306376A	CR	0.73	0.44	0
MA310860	SR	0.65	0.53	0
MA310891	CR	0.24	0.48	0
MA703074822	CR	0.34	0.59	1
MA703080328	CR	0.25	0.36	0
MA703131543	CR	0.32	0.63	1
MA713467960	CR	0.53	0.54	1
MA206980	SR	0.76	0.49	0
MA207009	SR	0.73	0.29	0
MA227228	SR	0.50	0.42	0
MA227291	CR	0.72	0.57	0
MA310890	CR	0.51	0.55	0
MA310895	CR	0.42	0.39	0
MA260962	SR	0.63	0.40	0
MA287138	SR	0.73	0.49	0
MA300056	CR	0.32	0.47	0
MA623065846	CR	0.70	0.54	0
MA203640A	CR	0.45	0.53	0
MA281992	SR	0.83	0.40	0
MA300749A	CR	0.56	0.49	0
MA306300	CR	0.62	0.51	0
MA309916A	CR	0.85	0.45	0
MA310842	SR	0.63	0.47	0
MA310856	SR	0.71	0.45	0
MA310869	SR	0.83	0.44	0
MA703056978	CR	0.62	0.50	0
MA703084007	CR	0.25	0.60	1
MA713745785	CR	0.55	0.43	0
MA724333304	CR	0.55	0.69	0
MA207001	SR	0.77	0.43	0
MA207791	SR	0.80	0.42	0
MA300732	SR	0.58	0.42	0
MA306339	CR	0.60	0.49	0

* blank values represent no omitted responses on an item, and 0% is a result of rounding for very small values.

Table E-8. Item-Level Classical Test Theory Statistics—Mathematics Grade 4

<i>Item</i>		<i>Difficulty</i>	<i>Discrimination</i>	<i>Percent Omitted (%)</i>
<i>Number</i>	<i>Type</i>			
MA227864	CR	0.78	0.30	0
MA311529	SR	0.47	0.38	*
MA623831709	CR	0.74	0.43	0
MA303329	SR	0.42	0.49	0
MA307033	CR	0.34	0.54	0
MA311551	SR	0.60	0.54	*
MA311583	CR	0.42	0.61	0
MA704646689	CR	0.38	0.56	0
MA704647848	CR	0.33	0.47	0
MA704650142	CR	0.64	0.42	0
MA713939739	CR	0.66	0.65	0
MA714225971	CR	0.45	0.52	1
MA714230904	CR	0.22	0.37	0
MA714233266	CR	0.69	0.58	0
MA222213	SR	0.63	0.11	*
MA227383	SR	0.77	0.41	0
MA227456	SR	0.86	0.34	0
MA247598	SR	0.92	0.30	0
MA298090	SR	0.39	0.33	0
MA303319	SR	0.78	0.50	0
MA311567	CR	0.57	0.52	0
MA279765	SR	0.55	0.57	*
MA306994	SR	0.41	0.44	0
MA307317	CR	0.50	0.60	0
MA311552	SR	0.49	0.57	0
MA704650539	CR	0.58	0.51	0
MA704652242	CR	0.45	0.29	0
MA704653374	CR	0.61	0.43	0
MA713680384	CR	0.36	0.34	0
MA714111699	CR	0.43	0.40	0
MA247529	SR	0.87	0.41	0
MA247745	SR	0.54	0.44	0
MA279791	CR	0.47	0.52	0
MA286777	SR	0.71	0.45	0
MA293812	CR	0.42	0.62	0
MA297614	SR	0.63	0.39	0
MA297625	SR	0.95	0.29	*
MA304988	CR	0.48	0.46	0
MA307037	CR	0.92	0.31	0
MA307067	SR	0.71	0.53	0

* blank values represent no omitted responses on an item, and 0% is a result of rounding for very small values.

Table E-9. Item-Level Classical Test Theory Statistics—Mathematics Grade 5

<i>Item</i>		<i>Difficulty</i>	<i>Discrimination</i>	<i>Percent Omitted (%)</i>
<i>Number</i>	<i>Type</i>			
MA280507	SR	0.35	0.33	0
MA287252	SR	0.48	0.38	0
MA306425	CR	0.37	0.54	0
MA306465	SR	0.76	0.35	0
MA311301	SR	0.33	0.17	0
MA311329	CR	0.23	0.49	0
MA624344396	CR	0.58	0.51	0
MA306397	CR	0.36	0.48	0
MA306411	CR	0.26	0.44	0
MA306457	CR	0.52	0.68	0
MA715102122	CR	0.31	0.57	0
MA715102268	CR	0.60	0.45	0
MA715102367	CR	0.37	0.57	0
MA715102395	CR	0.32	0.51	0
MA715102462	CR	0.44	0.65	0
MA204866	SR	0.47	0.45	0
MA217315	SR	0.83	0.38	*
MA238611	CR	0.71	0.49	0
MA293830	SR	0.30	0.36	0
MA301147	SR	0.76	0.51	0
MA251317	SR	0.32	0.18	0
MA298031	SR	0.42	0.24	0
MA301589	SR	0.35	0.26	0
MA306441	SR	0.73	0.40	0
MA306448	SR	0.37	0.13	0
MA624358270	CR	0.44	0.51	0
MA280726	SR	0.64	0.22	*
MA311285	SR	0.46	0.33	0
MA311287	SR	0.68	0.43	*
MA704359215	CR	0.50	0.46	0
MA704359315	CR	0.23	0.35	0
MA704359624	CR	0.71	0.20	0
MA715102093	CR	0.87	0.24	0
MA715102228	CR	0.50	0.54	0
MA715102321	CR	0.32	0.62	0
MA204666	SR	0.66	0.46	*
MA272911	SR	0.70	0.52	*
MA280511	CR	0.39	0.60	0
MA303755	CR	0.51	0.41	0
MA303765	CR	0.49	0.65	0

* blank values represent no omitted responses on an item, and 0% is a result of rounding for very small values.

Table E-10. Item-Level Classical Test Theory Statistics—Mathematics Grade 6

<i>Item</i>		<i>Difficulty</i>	<i>Discrimination</i>	<i>Percent Omitted (%)</i>
<i>Number</i>	<i>Type</i>			
MA301226	SR	0.62	0.36	0
MA251432	CR	0.57	0.54	0
MA280690	SR	0.60	0.52	*
MA282127	SR	0.44	0.46	*
MA307273	SR	0.85	0.37	*
MA307338	SR	0.56	0.44	0
MA703149512	CR	0.64	0.40	0
MA703177181	CR	0.58	0.41	0
MA703178717	CR	0.24	0.43	0
MA703235915	CR	0.25	0.17	0
MA703251109	CR	0.61	0.63	1
MA713648266	CR	0.20	0.47	0
MA713663381	CR	0.33	0.25	0
MA713739169	CR	0.22	0.57	0
MA272283	SR	0.55	0.51	*
MA290253	CR	0.35	0.74	1
MA298171	SR	0.52	0.44	0
MA311664	CR	0.23	0.56	0
MA311703	SR	0.51	0.34	0
MA314807	SR	0.59	0.45	0
MA311661	SR	0.34	0.41	0
MA311688	SR	0.73	0.48	0
MA307217	SR	0.83	0.45	0
MA703148718	CR	0.68	0.45	0
MA703176270	CR	0.60	0.50	0
MA703179529	CR	0.14	0.21	0
MA703253363	CR	0.28	0.67	1
MA713677108	CR	0.25	0.38	0
MA713678325	CR	0.22	0.41	0
MA713935781	CR	0.20	0.65	1
MA714375741	CR	0.49	0.50	0
MA251299	SR	0.57	0.54	0
MA266127	SR	0.52	0.51	0
MA280684	SR	0.79	0.44	*
MA293846	SR	0.55	0.26	0
MA293850	SR	0.55	0.40	0
MA298159	SR	0.71	0.44	0
MA307294	SR	0.29	0.23	0
MA307345	SR	0.30	0.35	0
MA311659	CR	0.77	0.40	0

* blank values represent no omitted responses on an item, and 0% is a result of rounding for very small values.

Table E-11. Item-Level Classical Test Theory Statistics—Mathematics Grade 7

<i>Item</i>		<i>Difficulty</i>	<i>Discrimination</i>	<i>Percent Omitted (%)</i>
<i>Number</i>	<i>Type</i>			
MA306624	SR	0.27	0.05	0
MA311091	SR	0.41	0.23	0
MA624047703	CR	0.31	0.49	0
MA624055700	CR	0.35	0.59	0
MA703873828	CR	0.37	0.49	0
MA703943771	CR	0.26	0.69	1
MA713847917	CR	0.24	0.41	0
MA713848070	CR	0.27	0.55	0
MA713848101	CR	0.23	0.38	0
MA208761	CR	0.37	0.41	0
MA272149	SR	0.42	0.50	0
MA298218	SR	0.49	0.37	0
MA302315	SR	0.69	0.44	*
MA302320	SR	0.73	0.06	0
MA302339	CR	0.54	0.69	1
MA303730	SR	0.42	0.31	0
MA306604	SR	0.49	0.48	0
MA306606	CR	0.52	0.42	0
MA309787	SR	0.33	0.16	0
MA311107	SR	0.39	0.21	0
MA306625	CR	0.12	0.48	0
MA311110	CR	0.10	0.30	0
MA208422	SR	0.44	0.43	0
MA228036	SR	0.52	0.31	0
MA298068	SR	0.20	0.32	0
MA301853	CR	0.17	0.54	0
MA703880042	CR	0.23	0.26	0
MA703932370	CR	0.44	0.48	0
MA713847985	CR	0.21	0.51	0
MA713849144	CR	0.20	0.57	0
MA715009326	CR	0.46	0.75	1
MA208657	SR	0.69	0.45	0
MA281688	SR	0.69	0.47	0
MA281696	CR	0.72	0.49	0
MA298187	SR	0.45	0.38	0
MA298192	CR	0.16	0.49	0
MA306566	CR	0.22	0.72	3
MA311092	CR	0.26	0.65	0
MA311145	CR	0.11	0.49	1
MA314790	CR	0.17	0.53	0

* blank values represent no omitted responses on an item, and 0% is a result of rounding for very small values.

Table E-12. Item-Level Classical Test Theory Statistics—Mathematics Grade 8

<i>Item</i>		<i>Difficulty</i>	<i>Discrimination</i>	<i>Percent Omitted (%)</i>
<i>Number</i>	<i>Type</i>			
MA311423	SR	0.38	0.31	0
MA301472	SR	0.47	0.45	0
MA307397	SR	0.31	0.24	0
MA704844913	CR	0.67	0.56	0
MA704872840	CR	0.52	0.77	2
MA715919507	CR	0.46	0.50	0
MA715919547	CR	0.29	0.54	0
MA715919730	CR	0.29	0.32	1
MA253802	CR	0.50	0.57	1
MA259251	SR	0.41	0.43	0
MA272928	SR	0.55	0.58	0
MA287535	SR	0.48	0.19	0
MA289831	SR	0.62	0.39	0
MA296084	SR	0.52	0.44	0
MA297524	SR	0.33	0.50	0
MA301471	SR	0.45	0.59	0
MA301485	CR	0.37	0.78	1
MA307398	CR	0.14	0.33	1
MA307544	SR	0.78	0.37	0
MA311419	CR	0.20	0.54	0
MA228155	SR	0.55	0.36	0
MA228156	SR	0.66	0.40	0
MA253756	SR	0.65	0.38	0
MA704834722	CR	0.25	0.23	1
MA704836757	CR	0.60	0.40	0
MA704848184	CR	0.24	0.46	0
MA715919577	CR	0.43	0.42	0
MA715919661	CR	0.61	0.23	0
MA715919810	CR	0.51	0.58	0
MA715919824	CR	0.56	0.42	0
MA275024	SR	0.44	0.45	0
MA275045	SR	0.64	0.46	0
MA287541	SR	0.63	0.41	0
MA287552	SR	0.75	0.42	0
MA287556	CR	0.31	0.44	0
MA297660	SR	0.36	0.42	0
MA301677	SR	0.37	0.48	0
MA301680	CR	0.28	0.69	2
MA307542	SR	0.29	0.16	0
MA311433	CR	0.34	0.69	2

APPENDIX F

SCORE DISTRIBUTIONS

Table F-1. Item-Level Score Distributions for SR and OR Items and WPs—ELA

Grade	Item Number	Total Possible Points	Percent of Students at Score Point					
			0	1	2	3	4	5
3	EL293246	2	17.47	25.79	56.71			
	EL293259	2	1.50	11.91	85.64			
	EL293264#SCORE_TRAIT_Conv	3	19.05	36.16	25.08	18.89		
	EL293264#SCORE_TRAIT_Ideadev	4	19.61	34.31	25.87	14.98	4.43	
	EL624585614	2	12.32	38.02	49.49			
	EL62458579	2	23.48	23.12	53.33			
	EL715954244#SCORE_TRAIT_Conv	3	8.72	70.09	17.83	2.20		
	EL715954244#SCORE_TRAIT_Ideadev	4	11.15	67.90	17.72	1.76	0.32	
	EL708638799	2	35.43	19.51	45.00			
	EL708639108	2	30.00	26.70	43.11			
	EL708642952	3	29.35	28.66	36.76	4.37		
4	EL707249352	2	22.45	8.76	68.79			
	EL707253226	2	15.14	34.78	50.03			
	EL707254137#SCORE_TRAIT_Conv	3	21.89	53.10	20.35	4.17		
	EL707254137#SCORE_TRAIT_Ideadev	4	14.95	59.26	21.03	3.76	0.51	
	EL710438990#SCORE_TRAIT_Conv	3	28.85	44.63	19.45	6.57		
	EL710438990#SCORE_TRAIT_Ideadev	4	29.15	42.95	20.42	5.70	1.28	
	EL710453346	2	5.28	55.38	39.28			
	EL710740005	2	41.55	18.85	39.59			
	EL712435018	2	9.84	41.23	48.65			
	EL712438196	3	15.30	23.31	55.23	5.89		
EL712444620	2	29.72	16.85	53.40				
5	EL709062113	2	25.46	23.01	51.50			
	EL709062171	2	28.76	24.43	46.75			
	EL709062207#SCORE_TRAIT_Conv	3	22.01	49.41	20.77	7.31		
	EL709062207#SCORE_TRAIT_Ideadev	4	13.98	58.02	21.03	5.90	0.57	
	EL709229186#SCORE_TRAIT_Conv	3	9.05	54.82	20.01	15.79		
	EL709229186#SCORE_TRAIT_Ideadev	4	32.98	31.35	19.11	11.46	4.77	
	EL709237137	2	12.55	15.66	71.69			
	EL624182427#SCORE_TRAIT_Conv	3	15.75	44.45	28.15	11.30		
	EL624182427#SCORE_TRAIT_Ideadev	4	21.43	39.43	27.41	9.84	1.54	
EL733854336	2	49.15	13.62	37.22				
EL733940008	2	18.81	4.31	76.87				
6	EL707541984	2	68.80	5.40	25.79			
	EL707542115	2	17.38	60.48	22.07			
	EL710355409#SCORE_TRAIT_Conv	3	17.50	32.30	32.16	17.45		
	EL710355409#SCORE_TRAIT_Ideadev	5	12.63	33.87	34.92	14.03	3.40	0.55
	EL712756190#SCORE_TRAIT_Conv	3	24.98	35.89	28.02	10.19		
	EL712756190#SCORE_TRAIT_Ideadev	5	20.64	40.05	27.97	8.04	1.99	0.39
	EL720341012	2	17.89	24.71	57.39			
	EL720544333	2	35.80	17.39	46.74			
	EL707351199#SCORE_TRAIT_Conv	3	22.45	34.49	23.19	18.90		
	EL707351199#SCORE_TRAIT_Ideadev	5	26.80	29.58	24.59	13.22	3.84	1.00
	EL707643593	2	36.19	11.35	52.46			
EL710269684	2	24.17	10.09	65.75				

continued

Grade	Item Number	Total Possible Points	Percent of Students at Score Point					
			0	1	2	3	4	5
7	EL285797#SCORE_TRAIT_Conv	3	19.77	44.55	26.55	8.35		
	EL285797#SCORE_TRAIT_Ideadev	5	14.13	49.90	26.82	7.36	0.93	0.07
	EL707931187	2	29.08	37.05	33.84			
	EL707935717#SCORE_TRAIT_Conv	3	16.08	31.60	31.67	20.14		
	EL707935717#SCORE_TRAIT_Ideadev	5	10.62	36.25	31.10	16.74	3.76	1.01
	EL711859903	2	23.32	8.42	68.26			
	EL718059830	2	35.68	16.84	47.45			
	EL718150308	2	50.56	7.59	41.84			
	EL714343909#SCORE_TRAIT_Conv	3	33.82	30.70	17.59	17.01		
	EL714343909#SCORE_TRAIT_Ideadev	5	47.19	15.59	19.16	11.37	4.59	1.23
	EL719947920	2	50.90	7.59	41.48			
	EL723042397	2	21.10	11.10	67.73			
	8	EL708956273	2	51.80	14.33	33.85		
EL709072879		2	61.73	10.08	28.11			
EL709130565#SCORE_TRAIT_Conv		3	16.08	24.92	29.29	28.80		
EL709130565#SCORE_TRAIT_Ideadev		5	14.07	26.64	28.58	21.36	6.62	1.82
EL709184717#SCORE_TRAIT_Conv		3	17.98	27.15	30.13	23.80		
EL709184717#SCORE_TRAIT_Ideadev		5	7.71	33.67	31.64	18.87	5.84	1.32
EL713050542		2	26.51	34.37	39.10			
EL728531887		2	9.11	53.01	37.72			
EL714443470		2	48.18	17.97	33.83			
EL714447304		2	16.95	30.95	52.07			
EL714447652#SCORE_TRAIT_Conv		3	22.33	37.04	24.44	14.68		
EL714447652#SCORE_TRAIT_Ideadev		5	35.94	23.84	24.40	11.36	2.44	0.53

Table F-2 Item-Level Score Distributions for SR and OR Items—Mathematics

Grade	Item Number	Total Possible Points	Percent of Students at Score Point					
			0	1	2	3	4	5
3	MA300745	1	45.82	54.12				
	MA283013A	3	38.41	40.51	4.67	16.10		
	MA300747	1	35.58	64.29				
	MA306376A	1	26.65	73.01				
	MA310891	1	76.23	23.75				
	MA703074822	1	64.77	34.14				
	MA703080328	1	75.06	24.78				
	MA703131543	3	44.38	30.18	7.66	17.23		
	MA713467960	1	45.55	53.27				
	MA227291	1	28.14	71.79				
	MA310890	1	48.81	51.02				
	MA310895	1	57.69	42.19				
	MA300056	1	68.31	31.66				
	MA623065846	1	29.37	70.48				
	MA203640A	1	55.26	44.68				
	MA300749A	1	43.97	55.87				
	MA306300	1	38.19	61.77				
	MA309916A	1	15.35	84.55				
	MA703056978	1	37.26	62.36				
	MA703084007	3	48.22	35.13	8.30	7.78		
	MA713745785	1	45.30	54.53				
	MA724333304	3	22.84	20.52	23.64	32.85		
MA306339	1	39.46	60.41					
4	MA227864	1	22.03	77.93				
	MA623831709	1	26.35	73.61				
	MA307033	1	65.79	34.18				
	MA311583	4	8.05	50.58	16.51	15.31	9.41	
	MA704646689	1	61.84	38.10				
	MA704647848	1	66.49	33.26				
	MA704650142	1	36.14	63.82				
	MA713939739	4	11.05	9.53	16.97	29.99	32.17	
	MA714225971	1	54.65	44.75				
	MA714230904	1	77.97	21.90				
	MA714233266	2	17.62	26.85	55.45			
	MA311567	1	43.15	56.68				
	MA307317	4	13.97	12.29	38.91	27.07	7.60	
	MA704650539	1	42.11	57.79				
	MA704652242	1	54.47	45.39				
	MA704653374	2	18.50	41.29	40.03			
	MA713680384	1	63.56	36.40				
	MA714111699	1	56.67	43.30				
	MA279791	1	52.61	47.25				
	MA293812	4	16.61	20.91	47.02	7.58	7.70	
	MA304988	1	51.50	48.33				
	MA307037	1	7.84	92.15				

continued

Grade	Item Number	Total Possible Points	Percent of Students at Score Point					
			0	1	2	3	4	5
5	MA306425	1	62.52	37.34				
	MA311329	1	76.79	23.06				
	MA624344396	1	41.90	58.03				
	MA306397	1	63.49	36.38				
	MA306411	1	73.62	26.27				
	MA306457	4	24.07	16.78	12.89	17.34	28.79	
	MA715102122	1	69.16	30.71				
	MA715102268	1	40.14	59.84				
	MA715102367	1	63.19	36.69				
	MA715102395	2	52.31	32.15	15.43			
	MA715102462	4	16.60	21.51	40.39	9.66	11.58	
	MA238611	1	28.65	71.26				
	MA624358270	2	32.53	46.13	21.28			
	MA704359215	1	50.28	49.51				
	MA704359315	1	77.11	22.86				
	MA704359624	1	29.18	70.80				
	MA715102093	1	13.41	86.57				
	MA715102228	1	49.78	50.07				
	MA715102321	1	67.60	32.30				
	MA280511	4	23.53	25.89	26.58	18.52	5.31	
MA303755	1	48.72	51.21					
MA303765	4	24.62	13.37	26.03	10.88	24.91		
6	MA251432	1	42.63	57.28				
	MA703149512	1	35.95	64.00				
	MA703177181	1	42.06	57.91				
	MA703178717	1	75.69	24.27				
	MA703235915	2	59.63	30.94	9.40			
	MA703251109	4	12.30	12.16	16.13	37.48	21.30	
	MA713648266	1	79.59	20.29				
	MA713663381	1	67.13	32.78				
	MA713739169	1	77.80	21.99				
	MA290253	4	24.82	33.93	25.76	6.82	8.13	
	MA311664	1	77.09	22.61				
	MA703148718	1	31.53	68.21				
	MA703176270	1	40.14	59.53				
	MA703179529	1	85.33	14.34				
	MA703253363	4	46.84	21.51	10.38	11.88	8.60	
	MA713677108	1	74.79	25.06				
	MA713678325	1	77.46	22.45				
	MA713935781	4	60.98	15.34	9.25	8.16	5.49	
	MA714375741	2	22.72	55.89	21.32			
	MA311659	1	22.47	77.40				

continued

Grade	Item Number	Total Possible Points	Percent of Students at Score Point					
			0	1	2	3	4	5
7	MA624047703	1	68.44	31.48				
	MA624055700	2	41.51	47.79	10.64			
	MA703873828	1	63.15	36.79				
	MA703943771	4	50.95	16.84	10.94	13.02	7.02	
	MA713847917	1	75.75	24.05				
	MA713848070	1	72.63	27.26				
	MA713848101	1	76.48	23.31				
	MA208761	1	62.45	37.47				
	MA302339	4	14.74	15.99	24.63	24.90	19.12	
	MA306606	1	47.99	51.90				
	MA306625	1	87.58	12.11				
	MA311110	1	89.29	10.45				
	MA301853	1	82.90	16.71				
	MA703880042	1	76.64	23.19				
	MA703932370	2	37.49	36.63	25.82			
	MA713847985	1	79.23	20.65				
	MA713849144	1	79.35	20.31				
	MA715009326	4	17.43	21.42	26.86	25.71	8.02	
	MA281696	1	28.20	71.62				
	MA298192	1	83.81	15.98				
MA306566	4	54.01	17.17	10.13	11.05	4.42		
MA311092	1	73.22	26.29					
MA311145	1	88.24	11.22					
MA314790	1	82.31	17.39					
8	MA704844913	2	9.83	45.94	44.07			
	MA704872840	4	19.94	22.06	13.51	11.83	30.74	
	MA715919507	1	53.39	46.35				
	MA715919547	1	70.68	29.22				
	MA715919730	1	70.00	29.02				
	MA253802	1	49.71	49.78				
	MA301485	4	18.36	42.52	19.34	8.63	9.90	
	MA307398	1	85.63	13.73				
	MA311419	1	79.90	19.75				
	MA704834722	1	74.47	24.74				
	MA704836757	1	39.77	60.05				
	MA704848184	2	61.51	28.69	9.57			
	MA715919577	1	56.96	42.93				
	MA715919661	1	38.78	61.14				
	MA715919810	1	48.78	51.11				
	MA715919824	1	43.43	56.45				
	MA287556	1	68.69	30.84				
	MA301680	4	27.12	43.30	16.98	7.66	2.59	
	MA311433	4	22.21	37.48	21.32	11.68	5.22	

APPENDIX G

DIFFERENTIAL ITEM FUNCTIONING RESULTS

Table G-1. Number of Items Classified as “Low” or “High” DIF, Overall and by Group Favored—ELA

Grade	Group		Item Type	Number of Items	Number “Low”			Number “High”			
	Reference	Focal			Total	Favoring		Total	Favoring		
						Reference	Focal		Reference	Focal	
3	Male	Female	SR	15	4	3	1	0	0	0	
			CR	7	0	0	0	0	0	0	
			ES	4	0	0	0	0	0	0	
	Not ELL	ELL	SR	15	5	5	0	1	1	0	
			CR	7	0	0	0	0	0	0	
			ES	4	0	0	0	0	0	0	
	Not Economically Disadvantaged	Economically Disadvantaged	SR	15	2	2	0	0	0	0	
			CR	7	0	0	0	0	0	0	
			ES	4	0	0	0	0	0	0	
	White	African American	SR	15	3	3	0	0	0	0	
			CR	7	1	1	0	0	0	0	
			ES	4	0	0	0	0	0	0	
		Hispanic / Latino	SR	15	1	1	0	0	0	0	
			CR	7	0	0	0	0	0	0	
			ES	4	0	0	0	0	0	0	
	Students Without Disabilities	Students with Disabilities	SR	15	1	1	0	0	0	0	
			CR	7	0	0	0	0	0	0	
			ES	4	1	1	0	0	0	0	
	4	Male	Female	SR	15	0	0	0	0	0	0
				CR	7	0	0	0	0	0	0
				ES	4	1	0	1	0	0	0
Not ELL		ELL	SR	15	4	4	0	1	1	0	
			CR	7	0	0	0	1	1	0	
			ES	4	0	0	0	0	0	0	
Not Economically Disadvantaged		Economically Disadvantaged	SR	15	1	1	0	0	0	0	
			CR	7	1	1	0	0	0	0	
			ES	4	0	0	0	0	0	0	
White		African American	SR	15	2	2	0	0	0	0	
			CR	7	1	1	0	0	0	0	
			ES	4	0	0	0	0	0	0	
		Hispanic / Latino	SR	15	3	3	0	0	0	0	
			CR	7	1	1	0	0	0	0	
			ES	4	0	0	0	0	0	0	

continued

Grade	Group		Item Type	Number of Items	Number "Low"			Number "High"			
	Reference	Focal			Total	Favoring		Total	Favoring		
						Reference	Focal		Reference	Focal	
4	Students Without Disabilities	Students with Disabilities	SR	15	2	2	0	0	0	0	
			CR	7	2	2	0	0	0	0	
			ES	4	2	2	0	0	0	0	
	Male	Female	SR	17	0	0	0	0	0	0	
			CR	5	0	0	0	0	0	0	
			ES	6	0	0	0	0	0	0	
	Not ELL	ELL	SR	17	6	6	0	0	0	0	
			CR	5	0	0	0	0	0	0	
			ES	6	0	0	0	0	0	0	
	5	Not Economically Disadvantaged	Economically Disadvantaged	SR	17	1	1	0	0	0	0
				CR	5	0	0	0	0	0	0
				ES	6	0	0	0	0	0	0
White		African American	SR	17	1	1	0	0	0	0	
			CR	5	2	2	0	0	0	0	
			ES	6	0	0	0	0	0	0	
White		Hispanic / Latino	SR	17	2	2	0	0	0	0	
			CR	5	0	0	0	0	0	0	
			ES	6	0	0	0	0	0	0	
6		Students Without Disabilities	Students with Disabilities	SR	17	3	3	0	0	0	0
				CR	5	0	0	0	0	0	0
				ES	6	0	0	0	0	0	0
	Male	Female	SR	15	2	1	1	0	0	0	
			CR	6	1	0	1	1	1	0	
			ES	6	1	0	1	0	0	0	
	Not ELL	ELL	SR	15	3	2	1	2	2	0	
			CR	6	1	1	0	0	0	0	
			ES	6	0	0	0	0	0	0	
	6	Not Economically Disadvantaged	Economically Disadvantaged	SR	15	3	3	0	0	0	0
				CR	6	1	1	0	0	0	0
				ES	6	0	0	0	0	0	0
White		African American	SR	15	2	0	2	2	2	0	
			CR	6	1	1	0	0	0	0	
			ES	6	0	0	0	0	0	0	
White		Hispanic / Latino	SR	15	3	3	0	0	0	0	
			CR	6	1	1	0	0	0	0	
			ES	6	0	0	0	0	0	0	

continued

Grade	Group		Item Type	Number of Items	Number "Low"			Number "High"			
	Reference	Focal			Total	Favoring		Total	Favoring		
						Reference	Focal		Reference	Focal	
6	Students Without Disabilities	Students with Disabilities	SR	15	2	2	0	0	0	0	
			CR	6	1	1	0	0	0	0	
			ES	6	2	2	0	0	0	0	
	Male	Female	SR	15	1	1	0	0	0	0	
			CR	6	1	1	0	0	0	0	
			ES	6	2	0	2	0	0	0	
	Not ELL	ELL	SR	15	1	1	0	1	1	0	
			CR	6	2	2	0	0	0	0	
			ES	6	0	0	0	0	0	0	
	7	Not Economically Disadvantaged	Economically Disadvantaged	SR	15	0	0	0	0	0	0
				CR	6	1	1	0	0	0	0
				ES	6	0	0	0	0	0	0
White		African American	SR	15	2	2	0	0	0	0	
			CR	6	1	1	0	0	0	0	
			ES	6	0	0	0	0	0	0	
White		Hispanic / Latino	SR	15	1	1	0	0	0	0	
			CR	6	0	0	0	0	0	0	
			ES	6	0	0	0	0	0	0	
8		Students Without Disabilities	Students with Disabilities	SR	15	2	1	1	0	0	0
				CR	6	0	0	0	0	0	0
				ES	6	3	3	0	0	0	0
	Male	Female	SR	15	1	1	0	0	0	0	
			CR	6	1	1	0	0	0	0	
			ES	6	1	0	1	0	0	0	
	Not ELL	ELL	SR	15	3	3	0	2	2	0	
			CR	6	0	0	0	0	0	0	
			ES	6	0	0	0	0	0	0	
	8	Not Economically Disadvantaged	Economically Disadvantaged	SR	15	1	1	0	0	0	0
				CR	6	0	0	0	0	0	0
				ES	6	0	0	0	0	0	0
White		African American	SR	15	2	2	0	0	0	0	
			CR	6	0	0	0	0	0	0	
			ES	6	0	0	0	0	0	0	
White		Hispanic / Latino	SR	15	2	2	0	0	0	0	
			CR	6	0	0	0	0	0	0	
			ES	6	0	0	0	0	0	0	

continued

Grade	Group		Item Type	Number of Items	Number "Low"			Number "High"		
	Reference	Focal			Total	Favoring		Total	Favoring	
						Reference	Focal		Reference	Focal
8	Students Without Disabilities	Students with Disabilities	SR	15	1	1	0	0	0	0
			CR	6	0	0	0	0	0	0
			ES	6	0	0	0	0	0	0

Table G-2. Number of Items Classified as "Low" or "High" DIF, Overall and by Group Favored—Mathematics

Grade	Group		Item Type	Number of Items	Number "Low"			Number "High"		
	Reference	Focal			Total	Favoring		Total	Favoring	
						Reference	Focal		Reference	Focal
3	Male	Female	SR	17	3	2	1	0	0	0
			CR	23	4	4	0	0	0	0
	Not ELL	ELL	SR	17	1	0	1	0	0	0
			CR	23	6	4	2	0	0	0
	Not Economically Disadvantaged	Economically Disadvantaged	SR	17	0	0	0	0	0	0
			CR	23	0	0	0	0	0	0
	White	African American	SR	17	3	2	1	0	0	0
			CR	23	5	2	3	1	1	0
		Hispanic / Latino	SR	17	1	0	1	0	0	0
			CR	23	1	1	0	0	0	0
	Students Without Disabilities	Students with Disabilities	SR	17	0	0	0	0	0	0
			CR	23	0	0	0	0	0	0
4	Male	Female	SR	18	1	1	0	0	0	0
			CR	22	3	3	0	0	0	0
	Not ELL	ELL	SR	18	0	0	0	0	0	0
			CR	22	3	2	1	0	0	0
	Not Economically Disadvantaged	Economically Disadvantaged	SR	18	0	0	0	0	0	0
			CR	22	0	0	0	0	0	0
	White	African American	SR	18	1	0	1	0	0	0
			CR	22	2	2	0	0	0	0
		Hispanic / Latino	SR	18	0	0	0	0	0	0
			CR	22	0	0	0	0	0	0
	Students Without Disabilities	Students with Disabilities	SR	18	5	3	2	0	0	0
			CR	22	7	6	1	1	1	0

continued

Grade	Group		Item Type	Number of Items	Number "Low"			Number "High"		
	Reference	Focal			Total	Favoring		Total	Favoring	
						Reference	Focal		Reference	Focal
5	Male	Female	SR	18	5	4	1	0	0	0
			CR	22	1	0	1	0	0	0
	Not ELL	ELL	SR	18	1	1	0	0	0	0
			CR	22	0	0	0	0	0	0
	Not Economically Disadvantaged	Economically Disadvantaged	SR	18	0	0	0	0	0	0
			CR	22	0	0	0	0	0	0
	White	African American	SR	18	3	2	1	0	0	0
			CR	22	4	3	1	0	0	0
			SR	18	0	0	0	0	0	0
			CR	22	0	0	0	0	0	0
	Students Without Disabilities	Students with Disabilities	SR	18	1	1	0	0	0	0
			CR	22	3	3	0	0	0	0
6	Male	Female	SR	20	5	4	1	0	0	0
			CR	20	2	0	2	0	0	0
	Not ELL	ELL	SR	20	1	1	0	0	0	0
			CR	20	0	0	0	0	0	0
	Not Economically Disadvantaged	Economically Disadvantaged	SR	20	1	1	0	0	0	0
			CR	20	0	0	0	0	0	0
	White	African American	SR	20	2	2	0	0	0	0
			CR	20	3	1	2	0	0	0
			SR	20	1	1	0	0	0	0
			CR	20	0	0	0	0	0	0
	Students Without Disabilities	Students with Disabilities	SR	20	1	1	0	1	1	0
			CR	20	7	6	1	0	0	0
7	Male	Female	SR	16	3	3	0	1	1	0
			CR	24	3	1	2	1	1	0
	Not ELL	ELL	SR	16	2	2	0	0	0	0
			CR	24	0	0	0	0	0	0
	Not Economically Disadvantaged	Economically Disadvantaged	SR	16	0	0	0	0	0	0
			CR	24	0	0	0	0	0	0
	White	African American	SR	16	3	3	0	0	0	0
			CR	24	1	1	0	0	0	0
			SR	16	2	1	1	0	0	0
			CR	24	0	0	0	0	0	0
	Students Without Disabilities	Students with Disabilities	SR	16	0	0	0	2	2	0
			CR	24	4	4	0	0	0	0

continued

Grade	Group		Item Type	Number of Items	Number "Low"			Number "High"		
	Reference	Focal			Total	Favoring		Total	Favoring	
						Reference	Focal		Reference	Focal
8	Male	Female	SR	21	0	0	0	0	0	0
			CR	19	1	1	0	0	0	0
	Not ELL	ELL	SR	21	1	0	1	0	0	0
			CR	19	3	2	1	0	0	0
	Not Economically Disadvantaged	Economically Disadvantaged	SR	21	0	0	0	0	0	0
			CR	19	0	0	0	0	0	0
	White	African American	SR	21	2	1	1	0	0	0
			CR	19	7	4	3	0	0	0
		Hispanic / Latino	SR	21	0	0	0	0	0	0
			CR	19	0	0	0	0	0	0
	Students Without Disabilities	Students with Disabilities	SR	21	5	5	0	0	0	0
			CR	19	2	1	1	0	0	0

APPENDIX H

RELIABILITY

Table H-1: Subgroup Reliabilities—ELA

Grade	Subgroup	Number of Students	Raw Score			Alpha	SEM
			Maximum	Mean	Standard		
3	All Students	10,188	44	23.34	8.20	0.87	2.90
	ELL	1,504	44	18.43	7.99	0.86	2.96
	Economically Disadvantaged	5,165	44	20.40	7.90	0.86	2.93
	African American	864	44	20.41	7.93	0.86	2.95
	Asian	350	44	25.33	7.70	0.86	2.87
	Hispanic	2,637	44	20.34	8.15	0.87	2.94
	Native American/Alaska Native	79	44	19.09	7.70	0.86	2.92
	White	5,711	44	25.17	7.70	0.86	2.85
	Pacific Islander/Hawaiian	20	44	21.80	8.26	0.88	2.89
	Multiracial	521	44	22.71	8.24	0.88	2.91
	Male	5,189	44	22.23	8.21	0.88	2.87
	Female	4,993	44	24.49	8.01	0.87	2.90
	Special Education	1,709	44	15.84	7.54	0.86	2.85
4	All Students	10,410	44	22.80	8.45	0.89	2.80
	ELL	1,564	44	17.56	7.91	0.87	2.87
	Economically Disadvantaged	5,282	44	19.70	7.97	0.87	2.85
	African American	910	44	19.77	8.18	0.88	2.86
	Asian	388	44	24.05	7.99	0.88	2.80
	Hispanic	2,732	44	19.68	8.08	0.87	2.87
	Native American/Alaska Native	85	44	18.38	8.46	0.89	2.84
	White	5,738	44	24.87	8.05	0.88	2.75
	Pacific Islander/Hawaiian	16	44	21.38	9.10	0.90	2.90
	Multiracial	527	44	21.54	8.44	0.89	2.80
	Male	5,349	44	22.06	8.45	0.89	2.77
	Female	5,047	44	23.58	8.39	0.89	2.82
	Special Education	1,511	44	13.49	6.67	0.83	2.71
5	All Students	10,694	48	25.28	9.22	0.89	3.08
	ELL	1,667	48	19.68	9.17	0.89	3.10
	Economically Disadvantaged	5,352	48	21.79	8.80	0.88	3.07
	African American	932	48	21.59	8.59	0.87	3.06
	Asian	391	48	27.50	8.88	0.88	3.11
	Hispanic	2,896	48	21.60	9.09	0.89	3.08
	Native American/Alaska Native	77	48	20.69	8.60	0.87	3.07
	White	5,895	48	27.68	8.61	0.87	3.05
	Pacific Islander/Hawaiian	9					
	Multiracial	494	48	24.34	8.79	0.88	3.07
	Male	5,400	48	24.08	9.23	0.89	3.04
	Female	5,294	48	26.51	9.06	0.88	3.10
	Special Education	1,481	48	15.24	7.44	0.84	2.93
6	All Students	10,792	51	23.62	10.69	0.91	3.21
	ELL	1,717	51	17.10	9.99	0.90	3.10
	Economically Disadvantaged	5,407	51	19.35	9.88	0.90	3.14
	African American	958	51	18.82	10.37	0.91	3.12
	Asian	322	51	26.76	10.40	0.90	3.23
	Hispanic	2,854	51	19.27	10.05	0.90	3.13

continued

Grade	Subgroup	Number of Students	Raw Score			Alpha	SEM
			Maximum	Mean	Standard		
6	Native American/Alaska Native	73	51	15.86	9.77	0.90	3.11
	White	6,032	51	26.39	10.02	0.90	3.20
	Pacific Islander/Hawaiian	20	51	21.05	10.50	0.91	3.16
	Multiracial	532	51	23.50	10.80	0.91	3.25
	Male	5,551	51	21.43	10.31	0.91	3.14
	Female	5,240	51	25.94	10.59	0.91	3.21
	Special Education	1,604	51	12.46	7.23	0.84	2.86
7	All Students	10,630	51	23.53	10.53	0.90	3.31
	ELL	1,584	51	16.66	9.28	0.88	3.18
	Economically Disadvantaged	5,056	51	18.95	9.22	0.88	3.21
	African American	854	51	18.72	9.21	0.88	3.18
	Asian	327	51	26.65	10.34	0.90	3.28
	Hispanic	2,783	51	18.73	9.38	0.88	3.20
	Native American/Alaska Native	65	51	17.23	9.52	0.89	3.13
	White	6,119	51	26.41	10.13	0.89	3.32
	Pacific Islander/Hawaiian	17	51	21.76	10.64	0.91	3.25
	Multiracial	444	51	21.44	10.09	0.89	3.28
	Male	5,503	51	21.55	10.31	0.90	3.23
	Female	5,106	51	25.63	10.34	0.90	3.34
	Special Education	1,514	51	13.30	7.21	0.83	2.97
	8	All Students	10,491	51	24.74	10.76	0.91
ELL		1,614	51	17.29	9.34	0.89	3.11
Economically Disadvantaged		4,827	51	19.95	9.65	0.89	3.18
African American		924	51	19.64	9.77	0.90	3.15
Asian		287	51	27.17	11.00	0.91	3.24
Hispanic		2,638	51	19.43	9.68	0.89	3.17
Native American/Alaska Native		79	51	18.42	8.29	0.85	3.17
White		6,122	51	27.91	10.10	0.89	3.28
Pacific Islander/Hawaiian		13	51	17.62	6.09	0.71	3.30
Multiracial		424	51	22.94	10.56	0.91	3.24
Male		5,308	51	22.52	10.46	0.90	3.25
Female		5,179	51	27.02	10.59	0.91	3.26
Special Education		1,510	51	14.15	7.38	0.84	2.91

Table H-2. Subgroup Reliabilities—Mathematics

Grade	Subgroup	Number of Students	Raw Score			Alpha	SEM
			Maximum	Mean	Standard		
3	All Students	10,321	48	24.10	11.69	0.93	3.00
	ELL	1,629	48	17.25	10.30	0.92	2.85
	Economically Disadvantaged	5,248	48	19.80	10.84	0.93	2.92
	African American	878	48	20.01	11.27	0.93	2.91
	Asian	360	48	28.47	11.43	0.93	2.97
	Hispanic	2,722	48	19.48	10.97	0.93	2.91
	Native American/Alaska Native	79	48	17.63	11.14	0.94	2.81
	White	5,732	48	26.82	11.16	0.93	3.03
	Pacific Islander/Hawaiian	20	48	20.50	11.66	0.94	2.95
	Multiracial	522	48	23.24	11.66	0.93	3.00
	Male	5,262	48	24.20	12.03	0.94	2.98
	Female	5,051	48	23.99	11.32	0.93	3.01
	Special Education	1,716	48	14.73	10.36	0.93	2.76
4	All Students	10,520	54	27.95	12.11	0.92	3.33
	ELL	1,672	54	21.29	11.72	0.92	3.31
	Economically Disadvantaged	5,351	54	23.65	11.36	0.92	3.31
	African American	917	54	23.37	11.48	0.92	3.27
	Asian	397	54	31.17	12.14	0.93	3.32
	Hispanic	2,810	54	23.62	11.71	0.92	3.33
	Native American/Alaska Native	84	54	23.57	13.12	0.94	3.33
	White	5,756	54	30.83	11.49	0.92	3.28
	Pacific Islander/Hawaiian	16	54	25.81	10.59	0.92	3.04
	Multiracial	526	54	25.85	11.97	0.92	3.34
	Male	5,409	54	28.45	12.58	0.93	3.34
	Female	5,097	54	27.42	11.59	0.92	3.31
	Special Education	1,507	54	15.33	9.29	0.89	3.10
5	All Students	10,807	54	23.79	12.05	0.91	3.59
	ELL	1,774	54	17.79	10.78	0.90	3.40
	Economically Disadvantaged	5,427	54	19.27	10.59	0.89	3.46
	African American	944	54	18.64	10.20	0.89	3.39
	Asian	398	54	29.50	11.86	0.91	3.56
	Hispanic	2,972	54	19.26	10.92	0.90	3.46
	Native American/Alaska Native	78	54	17.50	10.81	0.91	3.32
	White	5,911	54	26.71	11.87	0.91	3.60
	Pacific Islander/Hawaiian	9					
	Multiracial	495	54	22.56	11.51	0.90	3.55
	Male	5,461	54	23.81	12.62	0.92	3.58
	Female	5,346	54	23.78	11.45	0.90	3.59
	Special Education	1,481	54	12.36	7.88	0.85	3.01
6	All Students	10,912	54	22.82	11.68	0.92	3.38
	ELL	1,837	54	16.63	10.35	0.90	3.20
	Economically Disadvantaged	5,497	54	18.36	9.96	0.89	3.24
	African American	978	54	17.36	9.97	0.90	3.23
	Hispanic	2,942	54	18.12	10.15	0.90	3.24

continued

Grade	Subgroup	Number of Students	Raw Score			Alpha	SEM
			Maximum	Mean	Standard		
6	Native American/Alaska Native	74	54	14.96	8.58	0.87	3.08
	White	6,035	54	25.82	11.40	0.91	3.41
	Pacific Islander/Hawaiian	20	54	21.60	13.55	0.94	3.37
	Multiracial	535	54	22.26	12.13	0.92	3.37
	Male	5,612	54	22.40	11.88	0.92	3.37
	Female	5,299	54	23.26	11.46	0.91	3.38
	Special Education	1,603	54	12.02	7.10	0.83	2.92
7	All Students	10,735	54	18.96	11.24	0.91	3.37
	ELL	1,701	54	12.63	8.36	0.87	3.04
	Economically Disadvantaged	5,131	54	14.19	8.49	0.87	3.12
	African American	872	54	13.71	8.59	0.87	3.08
	Asian	335	54	24.06	12.80	0.93	3.50
	Hispanic	2,859	54	13.77	8.44	0.86	3.10
	Native American/Alaska Native	64	54	13.83	8.73	0.87	3.14
	White	6,129	54	22.03	11.46	0.91	3.43
	Pacific Islander/Hawaiian	17	54	15.35	8.89	0.87	3.15
	Multiracial	439	54	17.24	10.50	0.90	3.29
	Male	5,557	54	18.67	11.35	0.91	3.34
	Female	5,158	54	19.26	11.11	0.91	3.39
	Special Education	1,515	54	10.06	5.75	0.77	2.76
	8	All Students	10,610	54	22.61	11.87	0.92
ELL		1,734	54	15.57	9.48	0.89	3.12
Economically Disadvantaged		4,915	54	17.35	9.57	0.89	3.20
African American		932	54	17.45	9.88	0.90	3.19
Asian		290	54	27.13	13.12	0.93	3.37
Hispanic		2,727	54	16.70	9.60	0.89	3.18
Native American/Alaska Native		81	54	15.85	10.44	0.91	3.12
White		6,133	54	26.06	11.64	0.92	3.39
Pacific Islander/Hawaiian		14	54	14.21	6.83	0.81	2.95
Multiracial		430	54	20.50	11.50	0.92	3.35
Male		5,376	54	22.07	12.16	0.92	3.36
Female		5,231	54	23.16	11.53	0.91	3.38
Special Education		1,524	54	12.42	6.91	0.83	2.86

Table H-3. Reliabilities by Reporting Categories, Grade, and Content Area—ELA

Grade	Item Reporting Category	Label	Number of Items	Raw Score			Alpha	SEM
				Maximum	Mean	Standard Deviation		
3	1	Reading	19	27	16.35	5.47	0.81	2.39
	2	Language	5	9	4.43	1.99	0.60	1.26
	3	Writing	2	8	2.57	1.54	0.64	0.93
4	1	Reading	18	25	15.34	5.40	0.83	2.25
	2	Language	6	11	5.28	2.42	0.68	1.37
	3	Writing	2	8	2.19	1.45	0.69	0.81
5	1	Reading	18	23	14.91	5.12	0.81	2.26
	2	Language	7	13	6.67	2.74	0.70	1.50
	3	Writing	3	12	3.70	2.35	0.72	1.26
6	1	Reading	18	24	13.47	5.58	0.82	2.38
	2	Language	6	12	5.88	3.24	0.78	1.51
	3	Writing	3	15	4.27	2.79	0.84	1.12
7	1	Reading	18	24	13.95	5.60	0.81	2.46
	2	Language	6	12	5.49	3.13	0.76	1.55
	3	Writing	3	15	4.08	2.80	0.82	1.19
8	1	Reading	18	24	13.53	5.20	0.80	2.35
	2	Language	6	12	6.39	3.43	0.80	1.54
	3	Writing	3	15	4.82	3.09	0.87	1.13

Table H-4. Reliabilities by Reporting Categories, Grade, and Content Area—Mathematics

Grade	Item Reporting Category	Label	Number of Items	Raw Score			Alpha	SEM
				Maximum	Mean	Standard Deviation		
3	1	Operations and Algebraic Thinking	13	15	8.78	4.26	0.84	1.69
	2	Number and Operations in Base Ten	6	8	3.48	2.21	0.72	1.17
	3	Number and Operations-Fractions	7	9	4.34	2.55	0.74	1.30
	4	Measurement and Data	10	12	5.72	2.99	0.75	1.48
	5	Geometry	4	4	1.78	1.24	0.52	0.86
4	1	Operations and Algebraic Thinking	7	11	6.43	2.99	0.68	1.69
	2	Number and Operations in Base Ten	8	11	5.72	2.70	0.70	1.47
	3	Number and Operations-Fractions	13	16	8.41	4.09	0.84	1.64
	4	Measurement and Data	7	10	4.30	2.38	0.61	1.48
	5	Geometry	5	6	3.09	1.65	0.58	1.06
5	1	Operations and Algebraic Thinking	5	8	3.31	2.12	0.57	1.39
	2	Number and Operations in Base Ten	14	17	8.11	4.08	0.81	1.79
	3	Number and Operations-Fractions	9	13	4.88	3.15	0.64	1.89
	4	Measurement and Data	7	10	4.28	2.89	0.63	1.77
	5	Geometry	5	6	3.21	1.56	0.53	1.07
6	1	Ratios and Proportional Relationships	8	11	5.68	2.80	0.73	1.46
	2	The Number System	10	11	4.83	2.77	0.76	1.35
	3	Expressions and Equations	13	16	8.39	4.14	0.79	1.89
	4	Geometry	5	8	2.34	2.12	0.54	1.44
	5	Statistics and Probability	4	8	1.58	1.82	0.48	1.32
7	1	Ratios and Proportional Relationships	8	11	4.62	2.78	0.66	1.63
	2	The Number System	9	10	4.18	2.45	0.69	1.37
	3	Expressions and Equations	11	14	5.00	3.40	0.78	1.61
	4	Geometry	4	8	2.48	1.86	0.43	1.40
	5	Statistics and Probability	8	11	2.67	2.38	0.56	1.58
8	1	Number System & Expressions/Equations	17	21	8.48	5.11	0.85	1.97
	2	Functions	7	11	4.60	2.71	0.53	1.86
	3	Geometry	13	16	7.37	3.84	0.79	1.77
	4	Statistics and Probability	3	6	2.16	1.43	0.48	1.03