New York State Testing Program 2021: English Language Arts and Mathematics Grades 3–8

Technical Report

Questar Assessment Inc.
2021

Developed and published under contract with the New York State Education Department by Questar Assessment Inc., 5550 Upper 147th Street West, Apple Valley, MN 55124.

Table of Contents

Section 1: Introduction and Overview 1
1.1. Introduction 1
1.2. Test Purpose 2
1.3. Expected Participants 2
1.4. Test Use and Decisions Based on Assessment 3
1.4.1. Scale Scores 3
1.4.2. Performance Level Cut Scores and Classification 3
1.5. Testing Accommodations 4
1.6. Test Transcriptions 4
1.7. Test Translations 5
1.8. Test Participation and Results 5
Section 2: Test Design and Development 6
2.1. Test Descriptions 6
2.1.1. ELA and Mathematics Tests 6
2.2. Test Configuration 6
2.2.1. Test Design 6
2.3 Proficiency and Performance Standards 6
Section 3: Validity 8
3.1. Construct (Internal Structure) Validity 8
Section 4: Test Administration and Scoring 9
4.1. Test Administration 9
Section 5: Operational Test Data Collection and Classical Analysis 10
Section 6: IRT Analyses 11
6.1. Pre-Equating 11
Section 7: Reliability and Standard Error of Measurement 14
7.1. Test Reliability 14
7.1.1. Test Statistics and Reliability for Total Test 14
7.1.2. Test Reliability for Subgroups 16
7.2. Standard Error of Measurement (SEM) 25
7.3. Performance Level Classification Consistency and Accuracy 26
7.3.1. Consistency 26
7.3.2. Accuracy 28
Section 8: References 30
Appendix A: ELA and Mathematics Test Configurations 32
Appendix B: Operational Item Maps 34
Appendix C: Raw-to-Scale Score Tables 56
Appendix D: Study of Operational Test Mode Comparability 76
Appendix E: Memo on Operational Test Mode Comparability 86

List of Tables[1]

Table 1.1. 2021 Grades 3–8 Test Forms Original Administration Dates 4
Table 1.2. Overall Number and Percent of Students Tested & Not Tested in ELA and Math Statewide Across Grades 3–8 5
Table 6.1. Operational Scaling Coefficients 11
Table 6.2. ELA Scale Score Ranges Associated with Each Performance Level 12
Table 6.3. Mathematics Scale Score Ranges Associated with Each Performance Level 13
Table 7.1. ELA Test Form Statistics* 15
Table 7.2. ELA Test Reliability and Standard Error of Measurement* 15
Table 7.3. Mathematics Test Form Statistics* 15
Table 7.4. Mathematics Test Reliability and Standard Error of Measurement* 16
Table 7.5. ELA Grade 3 Test Reliability by Subgroup* 16
Table 7.6. ELA Grade 4 Test Reliability by Subgroup* 17
Table 7.7. ELA Grade 5 Test Reliability by Subgroup* 18
Table 7.8. ELA Grade 6 Test Reliability by Subgroup* 19
Table 7.9. ELA Grade 7 Test Reliability by Subgroup* 19
Table 7.10. ELA Grade 8 Test Reliability by Subgroup* 20
Table 7.11. Mathematics Grade 3 Test Reliability by Subgroup* 21
Table 7.12. Mathematics Grade 4 Test Reliability by Subgroup* 22
Table 7.13. Mathematics Grade 5 Test Reliability by Subgroup* 22
Table 7.14. Mathematics Grade 6 Test Reliability by Subgroup* 23
Table 7.15. Mathematics Grade 7 Test Reliability by Subgroup* 24
Table 7.16. Mathematics Grade 8 Test Reliability by Subgroup* 25
Table 7.17. Decision Consistency (All Cuts)* 27
Table 7.18. Decision Consistency (Level III Cut)* 28
Table 7.19. Decision Agreement (Accuracy) Estimates* 29
Table A.1. ELA Test Configuration 32
Table A.2. Mathematics Test Configuration 32
Table B.1. ELA Grade 3 Operational Item Map 34
Table B.2. ELA Grade 4 Operational Item Map 35
Table B.3. ELA Grade 5 Operational Item Map 37
Table B.4. ELA Grade 6 Operational Item Map 38
Table B.5. ELA Grade 7 Operational Item Map 40
Table B.6. ELA Grade 8 Operational Item Map 42
Table B.7. Mathematics Grade 3 Operational Item Map 44
Table B.8. Mathematics Grade 4 Operational Item Map 45
Table B.9. Mathematics Grade 5 Operational Item Map 47

[1] * Based on Session 1 test questions from the original administration test data.

Table B.10. Mathematics Grade 6 Operational Item Map 49
Table B.11. Mathematics Grade 7 Operational Item Map 51
Table B.12. Mathematics Grade 8 Operational Item Map 53
Table C.1. PBT ELA Grade 3 RSSS Table 56
Table C.2. PBT ELA Grade 4 RSSS Table 56
Table C.3. PBT ELA Grade 5 RSSS Table 57
Table C.4. PBT ELA Grade 6 RSSS Table 58
Table C.5. PBT ELA Grade 7 RSSS Table 59
Table C.6. PBT ELA Grade 8 RSSS Table 60
Table C.7. PBT Mathematics Grade 3 RSSS Table 60
Table C.8. PBT Mathematics Grade 4 RSSS Table 61
Table C.9. PBT Mathematics Grade 5 RSSS Table 62
Table C.10. PBT Mathematics Grade 6 RSSS Table 63
Table C.11. PBT Mathematics Grade 7 RSSS Table 63
Table C.12. PBT Mathematics Grade 8 RSSS Table 64
Table C.13. CBT ELA Grade 3 RSSS Table 66
Table C.14. CBT ELA Grade 4 RSSS Table 66
Table C.15. CBT ELA Grade 5 RSSS Table 67
Table C.16. CBT ELA Grade 6 RSSS Table 68
Table C.17. CBT ELA Grade 7 RSSS Table 69
Table C.18. CBT ELA Grade 8 RSSS Table 70
Table C.19. CBT Mathematics Grade 3 RSSS Table 71
Table C.20. CBT Mathematics Grade 4 RSSS Table 71
Table C.21. CBT Mathematics Grade 5 RSSS Table 72
Table C.22. CBT Mathematics Grade 6 RSSS Table 73
Table C.23. CBT Mathematics Grade 7 RSSS Table 74
Table C.24. CBT Mathematics Grade 8 RSSS Table 75
Table D.1. Summary of Item Type by Session on 2021 Grades 3–8 ELA Test Forms 76
Table D.2. Summary of Item Type on 2021 Grades 3–8 Math Test Forms 76
Table D.3. Test-level Performance between Test Modes Before Matching – ELA 79
Table D.4. Test-level Performance between Test Modes After Matching – ELA 80
Table D.5. Test-level Performance between Test Modes Before Matching – Math 80
Table D.6. Test-level Performance between Test Modes After Matching – Math 81
Table D.7. CBT Scale Score Adjustments Comparison after Matching 84

Section 1: Introduction and Overview

1.1. Introduction

This technical report provides detailed information regarding administration of the New York State Testing Program (NYSTP) Grades 3–8 English Language Arts (ELA) and Mathematics 2021 Operational Tests. This report includes information about test administration, test content and test statistics, validity and reliability, scoring, and scaling.

During the 2020–2021 school year, many students in New York State (NYS) received entirely remote instruction due to the COVID-19 pandemic. Given the U.S. Department of Education's decision not to approve the waiver requested by the New York State Education Department (NYSED) to forego assessments in 2021, it was determined that schools would be required to administer only Session 1 of the Grades 3–8 ELA and Mathematics 2021 Operational Tests to students. Further, only student scores on the Session 1 test questions would contribute to students' total scores for the tests. In addition, NYSED reduced the length of the tests by removing the embedded multiple-choice field test questions, and the tests contained previously administered questions. The decision to use previously administered test questions was based on guidance from nationally recognized experts in the assessment field. Reusing test questions provided the benefit of having established scale scores and stable item parameters.

NYSED also determined that it was not possible to administer the Grades 3–8 ELA and Mathematics tests remotely and that the assessments would be *school-based* and *in-person*. As such, students who received entirely remote instruction did not have to come to school only for testing. Moreover, to accommodate the variation in the way instruction was provided to students across the state during the 2020–2021 school year, NYSED provided schools much more flexibility than had been necessary in past years in scheduling the Spring 2021 operational tests. The 2020–21 Elementary- and Intermediate-Level Testing Schedule was amended to reflect that schools were able to use the entire testing window to administer the tests to students, whether the testing was paper-based (PBT) or computer-based (CBT). The ELA testing window was Monday, April 19 to Thursday, April 29 and the mathematics testing window was Monday, May 3 to Friday, May 14 to administer the tests to students, whether administration took place via PBT or CBT.

The following guidance was available to schools for the Spring 2021 administration of these tests as follows:

- Schools were not expected to bring students into the building to participate in the tests if the students were receiving entirely remote instruction during the testing windows. However, to assist in scheduling, NYSED suggested that schools consider contacting parents of students receiving entirely remote instruction to advise them of the testing schedule and ascertain the parent's interest in having their child come to school to participate in the test. Schools that were entirely remote during the entire testing window were not expected to re-open for the sole purpose of administering the tests.

- For students who were unable to participate in the test (Session 1) due to receiving entirely remote instruction during the entire testing window, schools reported this by

marking the field “receiving entirely remote instruction” on the student’s answer sheet or in the Nextera computer-based testing (CBT) system.

- Schools were given the option to administer the tests to students over the full duration of the testing window to accommodate instructional schedules.

- Schools were allowed to administer the tests to students who were receiving hybrid instruction on the days that they were ordinarily attending school in person. Schools were not required to test the entire grade on the same school day.

- Public school districts were not required to have a uniform testing schedule for all of the schools in the district.

- Make-up testing at each school for students who were absent on the day of their test were allowed to begin on the school day immediately following the first day of the test administration.

- The operational test did not include any embedded field test questions. As a result, there was one form administered per grade and subject and all questions in Session 1 contributed to student scores.

1.2. Test Purpose

The 2021 Grades 3–8 ELA and Mathematics NYSTP was designed to measure student knowledge and skills as defined by grade-level New York State Learning Standards in ELA and Mathematics. The tests were designed to allow the classification of student proficiency into four performance levels (Level I, Level II, Level III, and Level IV). Likewise, the test provided opportunities for students at each of these performance levels to demonstrate their knowledge and skills in the Learning Standards.

1.3. Expected Participants

Students in New York State public school Grades 3, 4, 5, 6, 7, and 8 (and ungraded students of equivalent chronological ages) were the expected participants for the Grades 3–8 NYSTP. Religious and independent schools could participate in the testing program, but their participation was not mandatory. Public school students were required to take all State assessments administered at their grade level, except for a very small percentage of students with severe cognitive disabilities who took the New York State Alternate Assessment (NYSAA), and for 2021, students who received exclusively remote instruction during the entire testing window. Schools that were entirely remote during the entire testing window were not expected to re-open for the sole purpose of administering the tests. For more detail on this exemption, please refer to the *NYSTP Grades 3–8 English Language Arts and Mathematics Tests School Administrator’s Manual* (SAM), available online at http://www.nysed.gov/common/nysed/files/programs/state-assessment/ei-sam-21ac.pdf.

1.4. Test Use and Decisions Based on Assessment

The NYSTP Grades 3–8 ELA and Mathematics tests are used to measure the extent to which individual students achieve the New York State Learning Standards in ELA and Mathematics, respectively. The results from the 2021 Grades 3–8 ELA and Mathematics Tests were provided to schools and parents for diagnostic purposes only and were to be utilized at the teacher and student level only as a way to help shape students' individualized learning plans to best meet that student's needs. This allowed educators to not only see how individual students were doing but to assess what was working in the classroom — and what wasn't — and make adjustments. Results could also be used to help educators better understand their students' need to provide supports more effectively.

Several types of scores are available from the Grades 3–8 ELA and Mathematics tests, and they are discussed in this section.

1.4.1. *Scale Scores*

The scale scores are a quantification of the proficiency measured by the Grades 3–8 ELA and Mathematics tests at each grade level. Scale scores are comparable only within a given subject and grade. Scale scores are not comparable across grades or across subjects. The scale scores are reported at the individual student level. Detailed information on the derivation and properties of the scale scores, including the range of scale scores for each subject and grade, is provided in Section 5: IRT Analyses.

1.4.2. *Performance Level Cut Scores and Classification*

Student performance is classified as Level I, Level II, Level III, or Level IV for the Grades 3–8 ELA and Mathematics tests. The definitions of performance levels are as follows:

- **NYS Level I:** Students performing at this level are well below proficient in standards for their grade. They demonstrate limited knowledge, skills, and practices embodied by the New York State P–12 Learning Standards for English Language Arts/Literacy or Mathematics that are considered insufficient for the expectations at this grade.

- **NYS Level II:** Students performing at this level are below proficient in standards for their grade. They demonstrate knowledge, skills, and practices embodied by the New York State P–12 Learning Standards for English Language Arts/Literacy or Mathematics that are considered partial but insufficient for the expectations at this grade.

- **NYS Level III:** Students performing at this level are proficient in standards for their grade. They demonstrate knowledge, skills, and practices embodied by the New York State P–12 Learning Standards for English Language Arts/Literacy or Mathematics that are considered sufficient for the expectations at this grade.

- **NYS Level IV:** Students performing at this level excel in standards for their grade. They demonstrate knowledge, skills, and practices embodied by the New York State P–12 Learning Standards for English Language Arts/Literacy or Mathematics that are considered more than sufficient for the expectations at this grade.

The performance level cut scores used to distinguish between Levels I, II, III, and IV were originally established during the process of standard-setting in summer 2013. In July 2018, Questar hosted a standards review meeting to revisit and update the established cut scores given a test design change and a reduced test length in 2018. The original standard-setting process is described in detail in Section 8 and Appendix P in the *2013 Technical Report* (NYSED, 2013). The *2018 Standards Review Report* is available in Appendix T of the *2018 Technical Report*.

For 2021, the test length was shortened further with only the Session 1 questions of the tests required to contribute to students' total scores. The Session 1-only tests were reviewed, and it was determined that the content remained aligned to the current test blueprints. In addition, since all items were previously administered (either in the Spring 2018 or Spring 2019 Operational Test administrations – see Table 1.1), they had established scale scores and stable item parameters. Therefore, the same cut scores and associated performance levels could be applied to these shortened tests.

Table 1.1. 2021 Grades 3–8 Test Forms Original Administration Dates

Grade	ELA	Math
3	2019	2018
4	2018	2019
5	2019	2019
6	2018	2018
7	2018	2019
8	2019	2019

1.5. Testing Accommodations

In accordance with federal law under the Americans with Disabilities Act and the section Fairness in Testing and Test Use in the *Standards for Educational and Psychological Testing* (AERA, APA, and NCME, 2014), accommodations that do not alter the measurement of any construct being tested are allowed for test takers. The allowance is in accordance with a student's Individualized Education Program (IEP) or Section 504 Accommodation Plan (504 Plan). School principals are responsible for ensuring that proper accommodations are provided when necessary, and that staff providing accommodations are properly trained. Details on testing accommodations can be found in the aforementioned *2021 School Administrator's Manual (SAM)*.

1.6. Test Transcriptions

For visually impaired students, large-type and Braille editions of the test books are provided. In most cases, the students dictate and/or record their responses, the teachers transcribe student responses to the multiple-choice items onto scannable answer sheets, and the teachers transcribe the responses to the constructed-response items onto the regular test books. Some of the students who use large-type editions will fill in the answer sheets by themselves. The large-type editions are created by Questar Assessment Inc. and printed by SeaChange Print Innovations. SeeWriteHear, LLC, produces the Braille editions. SeeWriteHear employs certified Library of Congress Braille transcribers and delivers Braille in accordance with the Braille Authority of North America (BANA) standards. Camera-ready versions of the regular test books are provided

to the Braille vendor, which then produces the Braille editions. Proofs of the Braille editions are submitted to NYSED for review and approval prior to production.

1.7. Test Translations

The NYSTP Grades 3–8 Mathematics tests are translated into eight languages: Arabic, Bengali, Chinese (Simplified), Chinese (Traditional), Haitian-Creole, Korean, Russian, and Spanish. These tests are translated to provide students the opportunity to demonstrate mathematical proficiency independent of their command of the English language. Sample tests are available in each translated language at the following location: https://www.nysedregents.org/ei/translations.html.

English Language Learner (ELL) students taking the Grades 3–8 Mathematics tests may be provided with an oral translation of the test when a written translation is not available in the student's native language. The following testing accommodations are also made available to ELLs: separate testing location, bilingual dictionaries and glossaries, simultaneous use of English and alternative-language editions, oral translation for lower-incidence languages, and writing responses in the native language.

The NYSTP Grades 3–8 ELA tests are not translated into any other language because they are assessments of proficiency in English language arts. The following testing accommodations are made available to ELLs taking the ELA tests: separate testing location and bilingual dictionaries and glossaries.

1.8. Test Participation and Results

Due to the circumstances related to the pandemic, approximately four out of ten students participated in the Spring 2021 Grades 3–8 ELA and Mathematics tests. In ELA, 41.9% of students enrolled in grades 3–8 took the assessment, and in mathematics, 39.9% of enrolled students took the assessment. Table 1.2 lists the number and percent of students tested statewide across Grades 3–8. Further details on test participation can be found in the press release here: http://www.nysed.gov/news/2021/state-education-department-releases-spring-2021-grades-3-8-ela-and-math-assessment-data.

Table 1.2. Overall Number and Percent of Students Tested & Not Tested in ELA and Math Statewide Across Grades 3–8

Category	ELA # of Students	ELA % Statewide	Math # of Students	Math % Statewide
Total Enrollment	1,195,169		1,195,044	
Students Who Took the State Tests	500,415	41.9%	476,753	39.9%
Students Who Did Not Take the State Tests	694,754	58.1%	718,291	60.1%

In an ordinary year, the percentage of students who participate in the tests is typically over 80 percent. Therefore, the 2021 Grades 3–8 assessment results are not representative of the state's student population and should not be compared to that of previous school years, statewide, or among subgroups of students. As a result, NYSED is not making statewide comparisons of the data, and only district- and school-level results are posted on the Department's website: https://data.nysed.gov/lists.php?type=district.

Section 2: Test Design and Development

2.1. Test Descriptions

The Grades 3–8 ELA and Mathematics tests are criterion-referenced tests composed of multiple-choice (MC) and constructed-response (CR) test items based on the New York State P–12 Learning Standards. For 2021, the tests were composed solely of MC items (Session 1). A second session (Session 2), consisting of predominantly CR items, was made available to schools strictly for their optional administration to students and, if administered, student scores for Session 2 test items were to be used for local student assessment purposes only. The tests were administered in New York State classrooms from April to May of 2021.

2.1.1. *ELA and Mathematics Tests*

For 2021, NYSED made the decision to reuse previously administered operational forms for the Grades 3–8 ELA and Mathematics tests (see Table 1.1). The details of test development are available in Section 2: Test Design and Development of the technical report produced for the year the forms were originally administered. The *2018 Technical Report* and *2019 Technical Report* are linked on the NYSED Technical Information and Reports Web page here: http://www.nysed.gov/state-assessment/grades-3-8-technical-information-and-reports.

2.2. Test Configuration

2.2.1. *Test Design*

The 2021 Grades 3–8 ELA and Mathematics tests were composed solely of MC items (Session 1). A second session (Session 2), consisting of predominantly CR items, was made available to schools strictly for their optional administration to students and, if administered, student scores for Session 2 test items were to be used for local student assessment purposes only.

The tables in Appendix A provide information on the numbers and types of items in each session for the Grades 3–8 ELA and Mathematics tests. NYSED further reduced the length of the Session 1 tests by removing the embedded multiple-choice field test questions.

2.3 Proficiency and Performance Standards

In July 2018, a standards review meeting occurred in Albany where New York State educators went through a rigorous process, guided by the best practices indicated by this intensely studied process, to recommend updated performance standards. These recommendations were presented to the Commissioner, who, in turn, adopted the recommended standards set forth by the committees. For additional details on the standards review process, see Appendix T of the *2018 Technical Report*.

Each grade level has four performance levels. Three cut points demarcate the performance levels needed to demonstrate each ascending level of performance. Section 6.1 contains the detailed information related to the performance standards (see Table 6.2 and Table 6.3.)

Following NYSED's directive that only Session 1 would be administered to students and the determination regarding which previously administered forms would be used for each content area and grade level, Questar Assessment, Inc's grade-level leads (senior assessment specialists) assessed each form for blueprint coverage and variety of item formats. Specifically, the grade-

level leads computed the percentage of blueprint coverage in each reporting category if a student were to take only Session 1 of a form. Notes were also made regarding the change in number of points derived from multiple choice and constructed response items if only Session 1 was taken. Based on these reviews, blueprint coverage and point distribution were deemed sufficient for the purposes of the 2021 administration. Therefore, for 2021, the same performance standards were used and applied to the shortened previously administered test forms for ELA and Mathematics.

Section 3: Validity

This section presents the evidence gathered to support the intended uses and interpretations of scores for the NYSTP Grades 3–8 English Language Arts (ELA) and Mathematics tests. All the test forms in the 2021 administration were previously administered in either the Spring 2018 or Spring 2019 Operational Test administrations. Table 1.1 lists the forms and their original year of use. Reusing test questions provided the benefit of having established scale scores and stable item parameters. All tests are considered pre-equated. Results for validity evidence (i.e., content validity, unidimensionality, and detection of bias) are available in Section 3: Validity of the original-year technical reports which can be found here: http://www.nysed.gov/state-assessment/grades-3-8-technical-information-and-reports.

In addition, schools were required to administer only Session 1 of the Grades 3–8 ELA and Mathematics tests to students, and only student scores on the Session 1 test questions contributed to students' total scores for the tests. Test reliability analysis on Session 1-only test questions included all tested New York State students who received valid scores from the Spring 2018 and Spring 2019 Operational Tests when the test forms were originally administered.

3.1. Construct (Internal Structure) Validity

Construct validity (i.e., what scores mean and what kind of inferences they support) is often considered the most important type of test validity. Empirical studies of the internal structure of the test provide one type of evidence of construct validity. For example, high internal consistency constitutes evidence of validity. This is because high coefficients imply that the test items are measuring the same domain of skill and are reliable and consistent.

Reliability coefficients for the 2021 ELA and Mathematics Tests were calculated for the shortened, previously administered Session 1 only questions using total populations and subgroups of students who received valid scores from the Spring 2018 or Spring 2019 Operational Tests, when the test forms were originally administered. The results are presented in Section 7.1: Test Reliability. For the total population, the ELA reliability coefficients (Cronbach's alpha) ranged from 0.77 to 0.86 on the Session 1 questions only. For all subgroups, the reliability coefficients were greater than or equal to 0.51. For the total population, the Mathematics reliability coefficients (Cronbach's alpha) ranged from 0.84 to 0.89 on the Session 1 questions only. For all subgroups, the reliability coefficients were greater than or equal to 0.64. Overall, high internal consistency of the NYSTP Grades 3–8 ELA and Mathematics Session 1 tests provided sound evidence of construct validity.

Section 4: Test Administration and Scoring

This section provides summaries of New York State test administration and scoring procedures. For further information, refer to the aforementioned *School Administrator's Manual* and the *New York State Scoring Leader Handbook (2021)* located here: http://www.nysed.gov/common/nysed/files/programs/state-assessment/3-8-scoring-leader-handbook-2021.pdf

4.1. Test Administration

The NYSTP Grades 3–8 ELA and Mathematics Tests were administered to students in a paper-based (PBT) and computer-based (CBT) testing mode in 2021. The 2020-21 Elementary- and Intermediate-Level Testing Schedule was amended to reflect that schools were able to use the entire testing window to administer the tests to students, whether the testing was paper-based or computer-based. The ELA testing window was Monday, April 19 to Thursday, April 29 for ELA and the Mathematics testing window was Monday, May 3 to Friday, May 14. Schools were given the option to administer the tests to students over the full duration of the testing window to accommodate instructional schedules.

Section 5: Operational Test Data Collection and Classical Analysis

Summaries of NYSTP Grades 3–8 ELA and Mathematics operational tests data collection, classical data analysis and DIF analyses for the 2021 Grades 3–8 ELA and Mathematics Session 1 tests are available in Section 5: Operational Test Data Collection and Classical Analysis of the technical report produced for the year the forms were originally administered. The *2018 Technical Report* and *2019 Technical Report* are linked on the NYSED Technical Information and Reports Web page found here: http://www.nysed.gov/state-assessment/grades-3-8-technical-information-and-reports.

Section 6: IRT Analyses

This section summarizes Item Response Theory (IRT) analyses for NYSTP Grades 3–8 ELA and Mathematics operational tests used in the scoring procedure to determine a student's total scores for the tests. All the Spring 2021 test forms were previously administered in either 2018 or 2019 (see Table 1.1.), and IRT calibration and post-equating were performed after the test forms' original administrations. The results of the IRT analyses in detail (i.e., calibration, post-equating and scaling processes) are available in Section 6 of the original-year technical reports found here: http://www.nysed.gov/state-assessment/grades-3-8-technical-information-and-reports. All tests in 2021 are considered pre-equated. The raw-score-to-scale-score conversion tables for the 2021 administration were developed using the item parameters of the Session 1 questions previously established in 2018 and 2019. For detailed procedures, see the following sections.

6.1. Pre-Equating

During the 2021 operational test administration, because all the test forms were previously administered in 2018 or 2019, pre-equating was conducted prior to administration of the tests. Based on the sample invariant property of IRT, all the item parameter estimates were obtained from large, representative samples, and placed on the reference scale in their previous administrations through the post-equating procedure described in the original-year technical reports. Item parameters derived in such a manner can be used to create raw-score-to-scale-score conversion tables prior to test administration. Neither calibration nor scaling was implemented in the pre-equating process.

The scale score is the basic score for the NYSTP. Raw-score-to-scale-score (RSSS) conversion tables based on the total number correct are presented in Appendix C, Table C.1–C.12. The raw-score-to-scale-score conversion tables for the 2021 administration were established using the inverse test characteristic curve (TCC) approach, where each student receives a theta score estimate corresponding to their raw score. For raw scores below the chance level or the perfect score, the following adjustment and interpolation was conducted to derive the adjusted theta scores: At the lower end of the scale, for any theta estimates that were lower than –2.5, 0.25 was subtracted from the preceding adjusted theta value that was within the range. At the higher end of the scale, for any theta estimates that were higher than 3.0, 0.25 was added to the previous theta value that was within the range. Then the scaling constants M_1^S and M_2^S (Table 6.1) were applied to derive the scale score of each student from their theta score estimate as follows:

$$ScaleScore = (M_1^S \cdot \theta) + M_2^S ,$$

Table 6.1. Operational Scaling Coefficients

Grade	Slope (M_1^S)	Intercept (M_2^S)
ELA		
3	18.310914	600.340994
4	18.276716	600.101132
5	18.212931	600.127742
6	18.309278	600.006654
7	18.318571	600.223246

8	18.308395	600.129092
Mathematics		
3	18.635919	600.082128
4	18.485491	600.009369
5	18.404109	600.040856
6	18.191784	600.432302
7	18.559827	600.499091
8	18.115200	600.640639

The conditional standard error of measurement (CSEM) of a scale score indicates the precision with which the proficiency, θ, is estimated, and it is inversely related to the amount of information provided by the test at a given proficiency level. The CSEM of a scale score is estimated as follows:

$$CSEM(SS) = M_1^S \cdot \frac{1}{\sqrt{I(\hat{\theta})}}$$

where $\hat{\theta}$ is the proficiency level corresponding to scale score, *SS*, $I(\hat{\theta})$ is the amount of information provided by the test at $\hat{\theta}$, and M_1^S is the scaling constant defined in Table 6.1. The CSEM value varies across proficiency levels and is the highest at the extreme ends of the scale where the amount of test information is typically the lowest.

The final element of the raw-score-to-scale-score tables is the application of the performance level cut scores. Scale score cuts were set in the summer of 2018 through standard review and were applied to the scale scores for the tests in the following years through the equating procedures. In 2021 raw-score-to-scale-score tables were calculated using the IRT parameters on the reference scale; therefore, they were put on the same scale as the original test. Also, content coverage review revealed that, based on session 1 only, blueprint coverage and point distribution were deemed sufficient across reporting categories for the purposes of the 2021 administration. Therefore, the cut scores of each performance level remain the same as those based on the original test. See Table 6.2 and Table 6.3 for the scale score ranges for each performance level.

Table 6.2 and Table 6.3 present scale score ranges associated with each performance level for ELA and Mathematics, respectively.

Table 6.2. ELA Scale Score Ranges Associated with Each Performance Level

Grade	NYS Level 1	NYS Level 2	NYS Level 3	NYS Level 4
3	542-582	583-601	602-628	629-642
4	542-583	584-602	603-618	619-655
5	523-593	594-608	609-621	622-656
6	523-589	590-601	602-613	614-652
7	523-590	591-606	607-622	623-650
8	523-583	584-602	603-616	617-647

Table 6.3. Mathematics Scale Score Ranges Associated with Each Performance Level

Grade	NYS Level 1	NYS Level 2	NYS Level 3	NYS Level 4
3	541-586	587-599	600-614	615-637
4	536-587	588-601	602-613	614-634
5	541-591	592-603	604-615	616-634
6	546-591	592-603	604-615	616-636
7	529-592	593-605	606-617	618-637
8	544-595	596-609	610-621	622-641

A mode comparability study was completed to identify whether or not there were any differences in student performance that could be attributed to the mode of test administration (i.e. PBT versus CBT). The main inference to be drawn from the mode comparability study is whether scores that arise from students testing on paper or on computer are interchangeable. A propensity score matching approach was conducted to generate the PBT and CBT samples that were comparable on covariates that may affect student performance, aside from the test mode itself (e.g., gender, school-type, previous performance) for the 2018 and 2019 administrations when the test forms were originally administered. This process included the entire test (Sessions 1 and 2). For the 2021 mode comparability study, the same matched samples were used to evaluate the difference in students' test scores using only the Session 1 test questions. The difference in students' test scores were computed between the matched PBT and CBT samples to evaluate test-level mode comparability, and mode adjustments were made accordingly. Please see Appendix D (the mode comparability study summary) and Appendix E (the NYSED memorandum on the mode comparability results) for more details.

Section 7: Reliability and Standard Error of Measurement

This section presents specific information on various test reliability statistics and standard error of measurement (SEM), as well as the results from a study of performance level classification accuracy and consistency. All the test forms in the 2021 administration were previously administered in either the Spring 2018 or Spring 2019 Operational Test administrations. Table 1.1 lists the forms and their original year of use. In addition, schools were required to administer only Session 1 of the Grades 3–8 ELA and Mathematics tests to students, and only student scores on the Session 1 test questions contributed to students' total scores for the tests. Test reliability analysis on the Session 1-only questions included all tested New York State students who received valid scores from the Spring 2018 or Spring 2019 Operational Tests when the test forms were originally administered.

7.1. Test Reliability

Test reliability is directly related to score stability and standard error and, as such, is an essential element of fairness and validity. Test reliability can be directly measured with an alpha statistic, or the alpha statistic can be used to derive the SEM. For the Grades 3–8 ELA and Mathematics tests, Questar calculated two types of reliability statistics: Cronbach's alpha (Cronbach, 1951) and Feldt-Raju coefficient (Qualls, 1995). These two measures are appropriate for assessment of a test's internal consistency when a single test is administered to a group of examinees on one occasion. The reliability of the test is then estimated by considering how well the items that reflect the same construct yield similar results (or how consistent the results are for different items that reflect the same construct measured by the test). Both Cronbach's alpha and Feldt-Raju coefficient measures are appropriate for tests of multiple-item formats (MC and CR items). In 2021, only Session 1 questions counted towards student scores. To determine if the Session 1-only tests were reliable, reliability coefficients were calculated for the shortened test. These reliability coefficients were calculated using Session 1 data from the original administration, either 2018 or 2019 (see Table 7.2 and Table 7.4). The reliability coefficients for the entire testing population and subgroups from 2018 or 2019 are presented for Session 1, only.

7.1.1. *Test Statistics and Reliability for Total Test*

Table 7.1 and Table 7.3 present the Session 1 test statistics including raw-score (RS) means and raw-score standard deviations (SDs) for ELA and Mathematics, respectively. These statistics give the necessary context for Table 7.2 and Table 7.4, which present the case counts (N-Count), number of test items (# Items), Cronbach's alpha and associated SEM, and Feldt-Raju coefficient and associated SEM obtained for the total ELA and Mathematics tests. Reliability coefficients provide measures of internal consistency that range from zero to one. High reliability indicates that scores are consistent and not unduly influenced by random error. Overall test reliability is a very good indication of each test's internal consistency.

Grades 3–8 ELA reliability estimates (Cronbach's alpha and Feldt-Raju) ranged from 0.77 to 0.86. Grades 3–8 Mathematics reliability estimates (Cronbach's alpha and Feldt-Raju) ranged from 0.84 to 0.89. The reliabilities are similar across grades and slightly higher for the Mathematics tests than for the ELA tests. All reliabilities were at least 0.77 across all grades and both subjects, which is a good indication that the NYSTP 2021 Grades 3–8 ELA and Mathematics tests are acceptably reliable.

Table 7.1. ELA Test Form Statistics*

	Item-Level			Student-Level			
	p-value				Raw Score		
Grade	Mean	Min.	Max.	N-Count	Max.	Mean	SD
3	0.61	0.36	0.87	177,503	18	11.06	3.99
4	0.61	0.48	0.81	181,672	18	10.96	3.91
5	0.64	0.34	0.87	175,383	28	17.82	5.96
6	0.63	0.36	0.87	170,015	28	17.71	5.53
7	0.59	0.36	0.87	155,919	28	16.63	5.52
8	0.66	0.38	0.88	151,966	28	18.39	5.75

**Note*. Based on Session 1 test questions from the original administration test data.

Table 7.2. ELA Test Reliability and Standard Error of Measurement*

			Raw Score	Cronbach's Alpha		Feldt-Raju Coefficient	
Grade	N-Count	Items	Points	Est.	SEM	Est.	SEM
3	177,503	18	18	0.80	1.77	0.80	1.76
4	181,672	18	18	0.77	1.86	0.77	1.86
5	175,383	28	28	0.86	2.22	0.86	2.21
6	170,015	28	28	0.84	2.24	0.84	2.23
7	155,919	28	28	0.82	2.33	0.82	2.31
8	151,966	28	28	0.86	2.19	0.86	2.18

**Note*. Based on Session 1 test questions from the original administration test data.

Table 7.3. Mathematics Test Form Statistics*

	Item-Level			Student-Level			
	p-value				Raw Score		
Grade	Mean	Min.	Max.	N-Count	Max.	Mean	SD
3	0.65	0.30	0.88	176,663	19	12.33	4.34
4	0.63	0.40	0.84	182,333	23	14.49	5.62
5	0.62	0.35	0.86	170,958	23	14.23	5.63
6	0.58	0.28	0.81	164,429	24	13.88	6.17
7	0.58	0.36	0.84	159,915	26	15.10	6.45
8	0.54	0.38	0.70	107,089	26	14.00	6.25

**Note*. Based on Session 1 test questions from the original administration test data.

Table 7.4. Mathematics Test Reliability and Standard Error of Measurement*

Grade	N-Count	Items	Raw Score Points	Cronbach's Alpha		Feldt-Raju Coefficient	
				Est.	SEM	Est.	SEM
3	176,663	19	19	0.84	1.74	0.84	1.73
4	182,333	23	23	0.88	1.95	0.88	1.94
5	170,958	23	23	0.88	1.95	0.88	1.95
6	164,429	24	24	0.89	2.01	0.89	2.00
7	159,915	26	26	0.89	2.13	0.89	2.12
8	107,089	26	26	0.87	2.23	0.87	2.22

**Note*. Based on Session 1 test questions from the original administration test data.

7.1.2. *Test Reliability for Subgroups*

The reporting subgroups include the following: gender, ethnicity, NRC, ELL/MLL, all SWD, all SUA, SWD/SUA (includes examinees who are classified as having a disability and who use at least one disability-related accommodation), and ELL/MLLs using accommodations specific to their ELL/MLL status (ELL/MLL/SUA). Accommodations available to students include the following: Flexibility in Scheduling/Timing, Flexibility in Setting, Method of Presentation (excluding Braille), Method of Response, Braille and Large-type, and others. Accommodations available to ELL/MLLs are Separate Location, and Bilingual Dictionaries and Glossaries.

As shown in Tables 7.5–7.10 and Tables 7.11–7.16 for ELA and Mathematics, respectively, the estimated reliabilities for most subgroups were close in magnitude to the test reliability estimates of the population tested in either 2018 or 2019. Cronbach's alpha reliability coefficients were all at least 0.51. Feldt-Raju reliability coefficients, which tend to be slightly larger than the Cronbach's alpha estimates for the same group, were at least 0.52.

Table 7.5. ELA Grade 3 Test Reliability by Subgroup*

Demographic Category		N-Count	Cronbach's Alpha		Feldt-Raju Coefficient	
			Est.	SEM	Est.	SEM
State	All Items	177,503	0.80	1.77	0.80	1.76
Gender	Female	87,733	0.80	1.76	0.80	1.75
	Male	89,770	0.81	1.78	0.81	1.77
Ethnicity	Asian	17,600	0.80	1.70	0.80	1.69
	African American	30,276	0.80	1.81	0.80	1.80
	Hispanic	50,893	0.78	1.83	0.78	1.83
	American Indian	1,213	0.78	1.80	0.78	1.80
	Multiracial	5,514	0.82	1.72	0.82	1.72
	Pacific Islander	388	0.81	1.75	0.81	1.75
	White	71,495	0.80	1.73	0.80	1.72
NRC	New York	64,285	0.81	1.79	0.81	1.78
	Big 4 Cities	7,405	0.77	1.85	0.77	1.85

Demographic Category		N-Count	Cronbach's Alpha		Feldt-Raju Coefficient	
			Est.	SEM	Est.	SEM
	Urban/Suburban	14,459	0.77	1.83	0.77	1.83
	Rural	9,685	0.77	1.80	0.78	1.79
	Average Needs	42,930	0.78	1.76	0.78	1.75
	Low Needs	19,085	0.77	1.66	0.77	1.65
	Charter School	12,483	0.78	1.71	0.79	1.70
	Religious and Independent	7,171	0.83	1.77	0.83	1.76
SWD	All Codes	26,406	0.74	1.88	0.74	1.88
SUA	All Codes	23,348	0.73	1.89	0.73	1.88
ELL/MLL	ELL/MLL=Y	19,028	0.66	1.91	0.67	1.91
SWD/SUA	SWD & SUA codes	20,718	0.71	1.89	0.71	1.89
ELL/MLL/SUA	SUA & ELL/MLL codes	3,640	0.59	1.93	0.59	1.92

**Note*. Based on Session 1 test questions from the original administration test data.

Table 7.6. ELA Grade 4 Test Reliability by Subgroup*

Demographic Category		N-Count	Cronbach's Alpha		Feldt-Raju Coefficient	
			Est.	SEM	Est.	SEM
State	All Items	181,672	0.77	1.86	0.77	1.86
Gender	Female	89,676	0.77	1.85	0.77	1.85
	Male	91,996	0.77	1.87	0.77	1.87
Ethnicity	Asian	18,533	0.77	1.77	0.77	1.76
	African American	32,133	0.76	1.90	0.76	1.90
	Hispanic	50,017	0.75	1.91	0.75	1.91
	American Indian	1,258	0.74	1.92	0.74	1.91
	Multiracial	4,731	0.78	1.83	0.78	1.83
	Pacific Islander	501	0.79	1.83	0.79	1.83
Ethnicity	White	72,411	0.76	1.83	0.76	1.83
NRC	New York	66,945	0.78	1.86	0.78	1.86
	Big 4 Cities	7,754	0.73	1.95	0.73	1.94
	Urban/Suburban	13,395	0.74	1.93	0.75	1.93
	Rural	9,820	0.74	1.92	0.74	1.92
	Average Needs	40,780	0.74	1.87	0.74	1.86
	Low Needs	18,128	0.72	1.75	0.72	1.74
	Charter School	11,288	0.74	1.82	0.74	1.82
	Religious and Independent	13,562	0.78	1.85	0.78	1.85
SWD	All Codes	26,145	0.71	1.96	0.71	1.95
SUA	All Codes	25,266	0.70	1.96	0.70	1.96
ELL/MLL	ELL/MLL=Y	17,497	0.62	1.98	0.62	1.98

Demographic Category		N-Count	Cronbach's Alpha		Feldt-Raju Coefficient	
			Est.	SEM	Est.	SEM
SWD/SUA	SWD & SUA codes	21,075	0.69	1.96	0.69	1.96
ELL/MLL/SUA	SUA & ELL/MLL codes	3,692	0.56	1.97	0.56	1.97

**Note*. Based on Session 1 test questions from the original administration test data.

Table 7.7. ELA Grade 5 Test Reliability by Subgroup*

Demographic Category		N-Count	Cronbach's Alpha		Feldt-Raju Coefficient	
			Est.	SEM	Est.	SEM
State	All Items	175,383	0.86	2.22	0.86	2.21
Gender	Female	86,147	0.86	2.19	0.86	2.18
	Male	89,236	0.86	2.24	0.86	2.24
Ethnicity	Asian	18,488	0.85	2.06	0.85	2.05
	African American	31,361	0.85	2.30	0.85	2.29
	Hispanic	49,603	0.85	2.29	0.85	2.29
	American Indian	1,175	0.84	2.28	0.84	2.27
	Multiracial	4,768	0.87	2.18	0.87	2.17
	Pacific Islander	456	0.85	2.18	0.85	2.17
	White	69,342	0.86	2.17	0.86	2.16
NRC	New York	65,311	0.86	2.22	0.86	2.21
	Big 4 Cities	7,597	0.85	2.35	0.85	2.34
	Urban/Suburban	13,673	0.85	2.32	0.85	2.31
	Rural	9,640	0.84	2.30	0.84	2.29
	Average Needs	40,861	0.84	2.21	0.85	2.20
	Low Needs	18,487	0.83	2.04	0.83	2.04
	Charter School	12,070	0.85	2.19	0.85	2.18
	Religious and Independent	7,744	0.88	2.21	0.88	2.21
SWD	All Codes	28,249	0.81	2.39	0.82	2.39
SUA	All Codes	27,271	0.82	2.39	0.82	2.39
ELL/MLL	ELL/MLL=Y	15,110	0.75	2.42	0.75	2.41
SWD/SUA	SWD & SUA codes	23,858	0.80	2.40	0.80	2.40
ELL/MLL/SUA	SUA & ELL/MLL codes	3,658	0.70	2.42	0.70	2.42

**Note*. Based on Session 1 test questions from the original administration test data.

Table 7.8. ELA Grade 6 Test Reliability by Subgroup*

Demographic Category		N-Count	Cronbach's Alpha		Feldt-Raju Coefficient	
			Est.	SEM	Est.	SEM
State	All Items	170,015	0.84	2.24	0.84	2.23
Gender	Female	83,617	0.83	2.20	0.83	2.19
	Male	86,398	0.84	2.28	0.84	2.26
Ethnicity	Asian	18,003	0.83	2.12	0.83	2.11
	African American	31,314	0.82	2.33	0.82	2.31
	Hispanic	46,768	0.82	2.32	0.82	2.31
	American Indian	1,141	0.83	2.29	0.83	2.28
	Multiracial	3,714	0.83	2.18	0.84	2.17
	Pacific Islander	614	0.83	2.20	0.83	2.19
	White	66,335	0.83	2.17	0.83	2.16
NRC	New York	64,138	0.84	2.26	0.84	2.25
	Big 4 Cities	6,856	0.82	2.36	0.83	2.35
	Urban/Suburban	11,921	0.82	2.33	0.82	2.32
	Rural	8,994	0.81	2.30	0.81	2.28
	Average Needs	36,469	0.81	2.21	0.81	2.20
	Low Needs	17,522	0.78	2.06	0.78	2.05
	Charter School	11,389	0.81	2.25	0.81	2.24
	Religious and Independent	12,726	0.84	2.23	0.84	2.22
SWD	All Codes	25,249	0.77	2.40	0.78	2.39
SUA	All Codes	23,977	0.78	2.40	0.78	2.39
ELL/MLL	ELL/MLL=Y	13,503	0.71	2.43	0.71	2.42
SWD/SUA	SWD & SUA codes	19,801	0.76	2.41	0.76	2.40
ELL/MLL/SUA	SUA & ELL/MLL codes	3,081	0.66	2.42	0.66	2.41

**Note*. Based on Session 1 test questions from the original administration test data.

Table 7.9. ELA Grade 7 Test Reliability by Subgroup*

Demographic Category		N-Count	Cronbach's Alpha		Feldt-Raju Coefficient	
			Est.	SEM	Est.	SEM
State	All Items	155,919	0.82	2.33	0.82	2.31
Gender	Female	75,962	0.81	2.30	0.81	2.29
Gender	Male	79,957	0.83	2.34	0.83	2.33
Ethnicity	Asian	17,046	0.84	2.19	0.84	2.18
	African American	29,642	0.79	2.38	0.79	2.37
	Hispanic	42,405	0.79	2.38	0.79	2.37
	American Indian	1,209	0.81	2.36	0.81	2.34
	Multiracial	2,969	0.84	2.28	0.84	2.26
	Pacific Islander	461	0.82	2.31	0.82	2.29

Demographic Category		N-Count	Cronbach's Alpha		Feldt-Raju Coefficient	
			Est.	SEM	Est.	SEM
Ethnicity	White	60,427	0.82	2.28	0.82	2.27
NRC	New York	64,280	0.83	2.32	0.83	2.31
	Big 4 Cities	6,366	0.79	2.40	0.79	2.39
	Urban/Suburban	10,852	0.79	2.39	0.79	2.38
	Rural	8,368	0.79	2.36	0.79	2.36
	Average Needs	32,952	0.80	2.32	0.80	2.31
	Low Needs	17,060	0.79	2.21	0.79	2.20
	Charter School	10,518	0.79	2.33	0.79	2.31
	Religious and Independent	5,523	0.85	2.32	0.85	2.31
SWD	All Codes	24,303	0.73	2.42	0.73	2.42
SUA	All Codes	23,095	0.74	2.42	0.74	2.41
ELL/MLL	ELL/MLL=Y	11,401	0.61	2.42	0.61	2.41
SWD/SUA	SWD & SUA codes	19,200	0.72	2.43	0.72	2.42
ELL/MLL/SUA	SUA & ELL/MLL codes	2,547	0.51	2.41	0.52	2.41

**Note*. Based on Session 1 test questions from the original administration test data.

Table 7.10. ELA Grade 8 Test Reliability by Subgroup*

Demographic Category		N-Count	Cronbach's Alpha		Feldt-Raju Coefficient	
			Est.	SEM	Est.	SEM
State	All Items	151,966	0.86	2.19	0.86	2.18
Gender	Female	73,609	0.85	2.16	0.85	2.15
	Male	78,357	0.86	2.22	0.86	2.21
Ethnicity	Asian	17,154	0.86	2.02	0.87	2.01
	African American	29,488	0.84	2.28	0.84	2.27
	Hispanic	42,959	0.84	2.27	0.84	2.27
	American Indian	1,215	0.84	2.26	0.84	2.25
	Multiracial	2,884	0.87	2.13	0.87	2.12
	Pacific Islander	447	0.85	2.16	0.85	2.15
	White	57,722	0.85	2.12	0.85	2.12
NRC	New York	63,250	0.85	2.20	0.86	2.19
	Big 4 Cities	6,284	0.85	2.34	0.85	2.33
	Urban/Suburban	10,461	0.84	2.31	0.84	2.31
	Rural	8,213	0.84	2.25	0.84	2.24
	Average Needs	31,286	0.85	2.17	0.85	2.17
	Low Needs	15,688	0.83	2.00	0.83	1.99
	Charter School	9,988	0.81	2.17	0.81	2.16
	Religious and Independent	6,796	0.86	2.13	0.86	2.13

Demographic Category		N-Count	Cronbach's Alpha		Feldt-Raju Coefficient	
			Est.	SEM	Est.	SEM
SWD	All Codes	24,635	0.79	2.40	0.79	2.39
SUA	All Codes	23,561	0.80	2.38	0.80	2.37
ELL/MLL	ELL/MLL=Y	10,856	0.69	2.44	0.69	2.43
SWD/SUA	SWD & SUA codes	20,218	0.78	2.40	0.78	2.39
ELL/MLL/SUA	SUA & ELL/MLL codes	2,458	0.60	2.44	0.60	2.44

*_Note_. Based on Session 1 test questions from the original administration test data.

Table 7.11. Mathematics Grade 3 Test Reliability by Subgroup*

Demographic Category		N-Count	Cronbach's Alpha		Feldt-Raju Coefficient	
			Est.	SEM	Est.	SEM
State	All Items	176,663	0.84	1.74	0.84	1.73
Gender	Female	87,245	0.83	1.75	0.84	1.74
	Male	89,418	0.84	1.74	0.84	1.73
Ethnicity	Asian	18,053	0.82	1.55	0.82	1.53
	African American	30,369	0.85	1.79	0.85	1.77
	Hispanic	50,760	0.83	1.80	0.83	1.79
	American Indian	1,228	0.83	1.77	0.84	1.76
	Multiracial	5,002	0.84	1.73	0.84	1.72
	Pacific Islander	422	0.83	1.69	0.83	1.68
	White	69,246	0.81	1.72	0.81	1.71
NRC	New York	67,143	0.85	1.74	0.85	1.73
	Big 4 Cities	6,623	0.83	1.86	0.83	1.84
	Urban/Suburban	14,296	0.82	1.84	0.83	1.82
	Rural	9,898	0.82	1.82	0.82	1.81
	Average Needs	42,171	0.81	1.76	0.82	1.75
	Low Needs	18,175	0.78	1.64	0.79	1.63
	Charter School	11,797	0.84	1.56	0.84	1.54
	Religious and Independent	6,560	0.81	1.79	0.81	1.78
SWD	All Codes	23,959	0.83	1.87	0.83	1.86
SUA	All Codes	22,053	0.81	1.88	0.81	1.87
ELL/MLL	ELL/MLL=Y	22,056	0.81	1.86	0.82	1.85
SWD/SUA	SWD & SUA codes	18,686	0.81	1.89	0.81	1.87
ELL/MLL/SUA	SUA & ELL/MLL codes	3,990	0.80	1.89	0.80	1.88

*_Note_. Based on Session 1 test questions from the original administration test data.

Table 7.12. Mathematics Grade 4 Test Reliability by Subgroup*

Demographic Category		N-Count	Cronbach's Alpha		Feldt-Raju Coefficient	
			Est.	SEM	Est.	SEM
State	All Items	182,333	0.88	1.95	0.88	1.94
Gender	Female	89,939	0.87	1.96	0.87	1.95
	Male	92,394	0.89	1.94	0.89	1.93
Ethnicity	Asian	18,721	0.88	1.68	0.88	1.66
	African American	31,428	0.87	2.03	0.87	2.02
	Hispanic	51,770	0.86	2.03	0.86	2.02
	American Indian	1,209	0.87	1.99	0.87	1.97
	Multiracial	5,275	0.89	1.92	0.89	1.91
	Pacific Islander	440	0.88	1.92	0.88	1.90
	White	73,279	0.87	1.91	0.87	1.90
NRC	New York	66,458	0.88	1.94	0.89	1.93
	Big 4 Cities	7,844	0.85	2.08	0.85	2.07
	Urban/Suburban	14,218	0.86	2.05	0.86	2.04
	Rural	9,892	0.86	2.03	0.86	2.02
	Average Needs	42,304	0.86	1.96	0.87	1.95
	Low Needs	19,086	0.85	1.78	0.86	1.77
	Charter School	11,210	0.88	1.81	0.88	1.80
	Religious and Independent	11,321	0.86	2.01	0.86	2.00
SWD	All Codes	27,122	0.83	2.10	0.83	2.09
SUA	All Codes	24,505	0.82	2.11	0.82	2.10
ELL/MLL	ELL/MLL=Y	19,764	0.81	2.12	0.81	2.11
SWD/SUA	SWD & SUA codes	21,634	0.80	2.12	0.80	2.11
ELL/MLL/SUA	SUA & ELL/MLL codes	3,848	0.76	2.13	0.77	2.12

**Note*. Based on Session 1 test questions from the original administration test data.

Table 7.13. Mathematics Grade 5 Test Reliability by Subgroup*

Demographic Category		N-Count	Cronbach's Alpha		Feldt-Raju Coefficient	
			Est.	SEM	Est.	SEM
State	All Items	170,958	0.88	1.95	0.88	1.95
Gender	Female	83,825	0.87	1.96	0.87	1.96
	Male	87,133	0.89	1.94	0.89	1.93
Ethnicity	Asian	18,888	0.88	1.70	0.88	1.68
	African American	30,274	0.87	2.03	0.87	2.02
	Hispanic	48,722	0.86	2.03	0.86	2.02
	American Indian	1,172	0.87	1.98	0.87	1.97
	Multiracial	4,500	0.88	1.93	0.89	1.93
	Pacific Islander	463	0.88	1.89	0.88	1.88

Demographic Category		N-Count	Cronbach's Alpha		Feldt-Raju Coefficient	
			Est.	SEM	Est.	SEM
Ethnicity	White	66,702	0.87	1.92	0.87	1.92
NRC	New York	65,842	0.88	1.95	0.89	1.94
	Big 4 Cities	7,051	0.86	2.06	0.86	2.05
	Urban/Suburban	12,678	0.85	2.05	0.85	2.05
	Rural	8,917	0.85	2.03	0.85	2.02
	Average Needs	39,166	0.86	1.96	0.86	1.95
	Low Needs	18,301	0.85	1.81	0.86	1.79
	Charter School	11,632	0.88	1.85	0.89	1.83
	Religious and Independent	7,371	0.88	2.03	0.88	2.02
SWD	All Codes	22,674	0.84	2.08	0.84	2.08
SUA	All Codes	20,176	0.83	2.09	0.83	2.08
ELL/MLL	ELL/MLL=Y	16,494	0.81	2.11	0.81	2.10
SWD/SUA	SWD & SUA codes	17,468	0.82	2.10	0.82	2.09
ELL/MLL/SUA	SUA & ELL/MLL codes	3,018	0.78	2.10	0.79	2.09

**Note*. Based on Session 1 test questions from the original administration test data.

Table 7.14. Mathematics Grade 6 Test Reliability by Subgroup*

Demographic Category		N-Count	Cronbach's Alpha		Feldt-Raju Coefficient	
			Est.	SEM	Est.	SEM
State	All Items	164,429	0.89	2.01	0.89	2.00
Gender	Female	80,609	0.89	2.02	0.89	2.01
	Male	83,820	0.90	2.00	0.90	1.99
Ethnicity	Asian	18,177	0.90	1.77	0.90	1.75
	African American	30,094	0.87	2.07	0.87	2.07
	Hispanic	46,282	0.87	2.08	0.87	2.07
	American Indian	1,143	0.89	2.03	0.89	2.02
	Multiracial	3,622	0.90	1.98	0.90	1.97
	Pacific Islander	605	0.89	1.96	0.89	1.95
	White	63,244	0.88	1.98	0.88	1.97
NRC	New York	63,931	0.90	2.00	0.90	1.99
	Big 4 Cities	5,499	0.87	2.07	0.87	2.06
	Urban/Suburban	12,070	0.86	2.09	0.87	2.08
	Rural	8,795	0.87	2.08	0.87	2.07
	Average Needs	36,022	0.88	2.02	0.88	2.01
	Low Needs	17,313	0.87	1.87	0.88	1.85
	Charter School	11,117	0.89	1.95	0.90	1.94
	Religious and Independent	9,682	0.87	2.07	0.87	2.07

Demographic Category		N-Count	Cronbach's Alpha		Feldt-Raju Coefficient	
			Est.	SEM	Est.	SEM
SWD	All Codes	23,672	0.83	2.09	0.83	2.08
SUA	All Codes	22,922	0.83	2.10	0.83	2.09
ELL/MLL	ELL/MLL=Y	14,402	0.82	2.10	0.82	2.09
SWD/SUA	SWD & SUA codes	19,225	0.81	2.09	0.81	2.09
ELL/MLL/SUA	SUA & ELL/MLL codes	3,263	0.75	2.08	0.75	2.07

*_Note_. Based on Session 1 test questions from the original administration test data.

Table 7.15. Mathematics Grade 7 Test Reliability by Subgroup*

Demographic Category		N-Count	Cronbach's Alpha		Feldt-Raju Coefficient	
			Est.	SEM	Est.	SEM
State	All Items	159,915	0.89	2.13	0.89	2.12
Gender	Female	78,603	0.89	2.14	0.89	2.12
	Male	81,312	0.89	2.13	0.89	2.12
Ethnicity	Asian	17,973	0.91	1.86	0.91	1.84
	African American	28,755	0.86	2.22	0.86	2.21
	Hispanic	45,020	0.86	2.22	0.86	2.21
	American Indian	1,093	0.88	2.19	0.88	2.18
	Multiracial	3,458	0.90	2.09	0.90	2.08
	Pacific Islander	589	0.90	2.07	0.90	2.05
	White	62,843	0.88	2.10	0.88	2.09
NRC	New York	63,197	0.90	2.11	0.90	2.10
	Big 4 Cities	6,113	0.83	2.25	0.83	2.24
	Urban/Suburban	10,944	0.83	2.25	0.83	2.24
	Rural	8,204	0.85	2.22	0.85	2.21
	Average Needs	33,003	0.87	2.15	0.87	2.14
	Low Needs	17,170	0.88	1.98	0.88	1.96
	Charter School	10,440	0.89	2.09	0.89	2.07
	Religious and Independent	10,844	0.88	2.18	0.88	2.17
SWD	All Codes	20,961	0.80	2.26	0.80	2.26
SUA	All Codes	18,621	0.80	2.26	0.80	2.26
ELL/MLL	ELL/MLL=Y	13,210	0.78	2.26	0.78	2.26
SWD/SUA	SWD & SUA codes	15,859	0.77	2.27	0.77	2.26
ELL/MLL/SUA	SUA & ELL/MLL codes	2,329	0.64	2.26	0.64	2.26

*_Note_. Based on Session 1 test questions from the original administration test data.

Table 7.16. Mathematics Grade 8 Test Reliability by Subgroup*

Demographic Category		N-Count	Cronbach's Alpha		Feldt-Raju Coefficient	
			Est.	SEM	Est.	SEM
State	All Items	107,089	0.87	2.23	0.87	2.22
Gender	Female	50,691	0.87	2.22	0.87	2.21
	Male	56,398	0.87	2.23	0.87	2.23
Ethnicity	Asian	10,401	0.90	1.97	0.90	1.96
	African American	22,170	0.85	2.27	0.85	2.27
	Hispanic	32,923	0.85	2.27	0.86	2.26
	American Indian	881	0.87	2.24	0.87	2.23
	Multiracial	1,918	0.87	2.24	0.87	2.23
	Pacific Islander	337	0.90	2.14	0.90	2.13
	White	38,276	0.86	2.23	0.86	2.22
NRC	New York	47,991	0.89	2.19	0.89	2.19
	Big 4 Cities	4,993	0.83	2.26	0.84	2.26
	Urban/Suburban	7,265	0.76	2.33	0.76	2.33
	Rural	6,221	0.82	2.31	0.82	2.30
	Average Needs	20,471	0.82	2.29	0.83	2.28
	Low Needs	8,292	0.86	2.16	0.86	2.15
	Charter School	6,605	0.89	2.11	0.89	2.10
	Religious and Independent	5,251	0.88	2.20	0.88	2.20
SWD	All Codes	17,506	0.78	2.30	0.78	2.29
SUA	All Codes	15,436	0.79	2.29	0.80	2.29
ELL/MLL	ELL/MLL=Y	10,081	0.80	2.29	0.80	2.29
SWD/SUA	SWD & SUA codes	13,166	0.77	2.29	0.77	2.29
ELL/MLL/SUA	SUA & ELL/MLL codes	1,689	0.64	2.29	0.64	2.29

**Note*. Based on Session 1 test questions from the original administration test data.

7.2. Standard Error of Measurement (SEM)

Table 7.2 and Table 7.4 present the SEMs, as computed from Cronbach's alpha and the Feldt-Raju reliability statistics, for Session 1 of the ELA and Mathematics tests, respectively. The SEMs ranged from 1.73 to 2.33 across subjects, grades, and the two methods of estimation, which is reasonable and small. The SEMs are directly related to reliability; the higher the reliability, the lower the standard error. As discussed, the reliability of these tests is relatively high, so it was expected that the SEMs would be low.

The SEMs for the subpopulations, as computed from Cronbach's alpha and the Feldt-Raju reliability statistics, are presented in Tables 7.5–7.10 and Tables 7.11–7.16. The SEMs associated with all reliability estimates for all subjects, grades, methods of estimation, and subpopulations ranged from 1.53 to 2.44, which is acceptably close to those for the entire population. This narrow range indicates that for Session 1 only of the Grades 3–8 ELA and Mathematics tests, all students' test scores were reasonably reliable with minimal error.

7.3. Performance Level Classification Consistency and Accuracy

This subsection describes the analyses conducted to estimate performance level classification consistency and accuracy for the Grades 3–8 ELA and Mathematics tests, based on Session 1 test questions, only. Performance level classification consistency and accuracy were calculated using Session 1 data from the original administration, either 2018 or 2019 (see Tables 7.17–7.19.) This provides statistical information on the classification of students into the four performance categories. Classification consistency refers to the estimated degree of agreement between examinees' performance classification from two independent administrations of the same test (or from two parallel forms of the test). Because obtaining test scores from two independent administrations of New York State tests was not feasible due to item release after each administration, a psychometric model was used to obtain the estimated classification consistency indices, using test scores from a single administration. Classification accuracy can be defined as the agreement between the actual classifications using observed cut scores and true classifications based on known true cut scores (Livingston and Lewis, 1995).

In conjunction with measures of internal consistency, classification consistency is an important type of reliability and is particularly relevant to high-stakes tests. As a form of reliability, classification consistency represents how reliably students can be classified into performance categories.

Classification consistency is most relevant for students whose performance is near the proficiency cut score. For example, consider the cut score delineating Levels II and III, or simply the "Level III Cut." Students whose proficiency is far above or far below that cut score are unlikely to be misclassified because repeated administration of the test will nearly always result in the same classification. Examinees whose true scores are close to the cut score are a more serious concern. These students' true scores will likely lie within the SEM of the cut score. For this reason, the measurement error at the cut scores should be considered when evaluating the classification consistency of a test. Furthermore, the number of students near the cut scores should also be considered when evaluating classification consistency; these numbers show the number of students who are at risk of being misclassified. Classification consistency and accuracy were estimated using the IRT procedure suggested by Lee, Hanson, and Brennan (2002) and Wang, Kolen, and Harris (2000). Appendix P in the original-year technical reports includes a description of the calculations and procedure based on the paper by Lee et al. (2002).

7.3.1. *Consistency*

The results for classifying students into four performance levels are separated from the results based solely on the Level III cut. Table 6.17 and Table 6.18 include case counts (N-Count), classification consistency (Agreement), classification inconsistency (Inconsistency), and Cohen's kappa (Kappa). Consistency indicates the rate at which a second administration would yield the same performance category designation (or a different designation for the inconsistency rate). The agreement index is a sum of the diagonal element in the contingency table. Kappa is similar but corrects for chance agreement. The inconsistency index is equal to the "1 - agreement index."

Table 7.17 depicts the ELA and Mathematics consistency study results, based on the range of performance levels for all grades. For ELA, 52–61% of students were estimated to be classified consistently to one of the four performance categories with a hypothetical second administration.

Kappa—that corrects for chance agreement—ranged from 0.34 to 0.46. These are between "fair" and "moderate" agreement, as per Landis and Koch's (1977) rules of thumb for kappa. For Mathematics, 61–67% of students were estimated to be classified consistently to one of the four performance categories, and kappa ranged from 0.47 to 0.56. These are considered "moderate" agreement, by Landis and Koch's (1977) rules of thumb for the kappa statistic. As mentioned above and for all tests, there is an acceptable amount of measurement error that all scores contain. By random chance, students testing twice may be classified first, for example, as a Level III and second as a Level IV. This is expected to occur more often for students scoring around the selected cut score, and less often for students closer to the middle of the performance level (i.e., close to the mid-point of two adjacent cut scores).

Table 7.17. Decision Consistency (All Cuts)*

Grade	N-Count	Agreement	Inconsistency	Kappa
ELA				
3	177,503	59%	41%	0.40
4	181,672	52%	48%	0.34
5	175,383	61%	39%	0.46
6	170,015	58%	42%	0.44
7	155,919	58%	42%	0.43
8	151,966	60%	40%	0.46
Mathematics				
3	176,663	61%	39%	0.47
4	182,333	65%	35%	0.53
5	170,958	65%	35%	0.53
6	164,429	67%	33%	0.56
7	159,915	66%	34%	0.54
8	107,089	64%	36%	0.51

**Note*. Based on Session 1 test questions from the original administration test data.
Decision consistency was calculated for PBT students only as item parameters were disproportionally based on PBT.

Table 7.18 depicts the ELA and Mathematics consistency study results based on two performance levels (NYS Level II and NYS Level III) as defined by the Level III cut. For ELA, 78–84% of the classifications of individual students were estimated to remain stable with a second administration. Kappa coefficients for ELA classification consistency ranged from 0.56 to 0.67. These are between "moderate" and "substantial" agreement, as per Landis and Koch's (1977) rules of thumb for kappa. For Mathematics, 85–90% of the classifications were estimated consistently, and kappa coefficients ranged from 0.68 to 0.79. These statistics indicate at least "substantial" agreement (where kappa > 0.60), as per Landis and Koch's (1977) rules of thumb for kappa.

Table 7.18. Decision Consistency (Level III Cut)*

Grade	N-Count	Agreement	Inconsistency	Kappa
ELA				
3	177,503	81%	19%	0.60
4	181,672	78%	22%	0.56
5	175,383	84%	16%	0.66
6	170,015	83%	17%	0.66
7	155,919	83%	17%	0.65
8	151,966	83%	17%	0.67
Mathematics				
3	176,663	85%	15%	0.68
4	182,333	87%	13%	0.75
5	170,958	87%	13%	0.73
6	164,429	88%	12%	0.77
7	159,915	90%	10%	0.79
8	107,089	90%	10%	0.77

**Note.* Based on Session 1 test questions from the original administration test data.
Decision consistency was calculated for PBT students only as item parameters were disproportionally based on PBT.

7.3.2. *Accuracy*

Table 7.19 presents the results of classification accuracy for the Session 1 only ELA and Mathematics tests across all grades. Included in the table are case counts (N-Count) and classification accuracy (Accuracy) for all performance levels (All Cuts) and for the Level III cut score. By definition, accuracy associated with the Level III cut is at least as great as that with the entire set of cut scores because there are only two categories for the former, as opposed to the latter, which has four.

For ELA, the estimated accuracy rates indicate that the categorization of a student's observed performance is in agreement with the location of his or her underlying proficiency from 60% to 70% of the time across all performance levels and 84% to 88% of the time in regard to the Level III cut score. For mathematics, the estimated accuracy rates indicate that the categorization of a student's observed performance is in agreement with the location of his or her true proficiency from 66% to 75% of the time across all performance levels and 86% to 93% of the time in regard to the Level III cut score.

Table 7.19. Decision Agreement (Accuracy) Estimates*

Grade	N-Count	Accuracy	
		All Cuts	Level III Cut
ELA			
3	177,503	67%	85%
4	181,672	60%	84%
5	175,383	70%	88%
6	170,015	67%	88%
7	155,919	65%	86%
8	151,966	67%	86%
Mathematics			
3	176,663	66%	86%
4	182,333	74%	91%
5	170,958	72%	90%
6	164,429	75%	91%
7	159,915	75%	93%
8	107,089	72%	93%

**Note*. Based on Session 1 test questions from the original administration test data.
Decision agreement was calculated for PBT students only as item parameters were disproportionally based on PBT.

Section 8: References

American Educational Research Association, American Psychological Association, and National Council on Measurement in Education (2014). *Standards for Educational and Psychological Testing*. Washington, D.C.: American Educational Research Association.

Bock, R.D. (1972). Estimating item parameters and latent ability when responses are scored in two or more nominal categories. *Psychometrika* 37: 29–51.

Cai, L., Thissen, D. J., & du Toit, S. (2011). IRTPRO (Version 2.1). Skokie, IL: Scientific Software International, Inc.

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

Cronbach, L.J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika* 16: 297–334.

Dorans, N.J., A.P. Schmitt & C.A. Bleistein (1992). The standardization approach to assessing comprehensive differential item functioning. *Journal of Educational Measurement* 29: 309–319.

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

Jensen, A.R. (1980). *Bias in mental testing*. New York: Free Press.

Kaiser, H. F. (1960). The application of electronic computers to factor analysis. *Educational and Psychological Measurement, 20,* 141–151.

Kim, S. & M. J. Kolen (2004). *STUIRT: A computer program for scale transformation under unidimensional item response theory models*. Iowa City, IA: Iowa Testing Programs, The University of Iowa.

Kolen, M.J. & Z. Cui (2004). *POLYEQUATE*. Iowa City, IA: Center for Advanced Studies in Measurement and Assessment, The University of Iowa.

Landis, J. R. & G. G. Koch. (1977). The Measurement of Observer Agreement for Categorical Data. *Biometrics, 33*(1), 159–174.

Lee, W. C., B.A. Hanson & R.L. Brennan (2002). Estimating consistency and accuracy indices for multiple classifications. *Applied Psychological Measurement* 26:412–432.

Lee, W. C. (2008). *Classification consistency and accuracy for complex assessments using item response theory*. (CASMA Research Report No. 27). Iowa City, IA: Center for Advanced Studies in Measurement and Assessment, The University of Iowa.

Lee, W. C. & M. J. Kolen (2006, Revised 2008). IRT-CLASS (Version 2.0). Iowa City, IA: Center for Advanced Studies in Measurement and Assessment, The University of Iowa.

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

Lord, F.M. (1980). *Applications of Item Response Theory to Practical Testing Problems.* Hillsdale, NJ: Lawrence Erlbaum.

Lord, F.M. & M.R. Novick (1968). *Statistical Theories of Mental Test Scores.* Menlo Park, CA: Addison-Wesley.

Mehrens, W.A. & I.J. Lehmann (1991). *Measurement and Evaluation in Education and Psychology, 3rd ed*. New York: Holt, Rinehart, and Winston.

Muraki, E. (1992). A generalized partial credit model: Application of an EM algorithm. *Applied Psychological Measurement* 16: 159–176.

NYSED. (2013) New York State Testing Program 2013: English Language Arts and Mathematics Grades 3–8 Technical Report. Albany, NY: New York State Education Department (NYSED). Retrieved from: http://www.p12.nysed.gov/assessment/reports/2013/ela-math-tr13.pdf

Qualls, A.L. (1995). Estimating the reliability of a test containing multiple-item formats. *Applied Measurement in Education* 8: 111–120.

Reckase, M.D. (1979). Unifactor latent trait models applied to multifactor tests: results and implications. *Journal of Educational Statistics* 4: 207–230.

Sandoval, J.H. & M.P. Mille (1979) *Accuracy of judgments of WISC-R item difficulty for minority groups*. Paper presented at the annual meeting of the American Psychological Association, New York. August.

Sinharay, S. (2006) Bayesian item fit analysis for unidimensional item response theory models. *British Journal of Mathematical and Statistical Psychology*, 59: 429–449.

Stocking, M.L. & F.M. Lord (1983). Developing a common metric in item response theory. *Applied Psychological Measurement* 7: 201–210.

Wang, T.M., J. Kolen, & D.J. Harris (2000). Psychometric properties of scale scores and performance levels for performance assessment using polytomous IRT. *Journal of Educational Measurement* 37: 141–162.

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

Yen, W.M. (1984). Obtaining maximum likelihood trait estimates from number correct scores for the three-parameter logistic model. *Journal of Educational Measurement*21: 93–111.

Zwick, R., J.R. Donoghue & A. Grima, (1993). Assessment of differential item functioning for performance tasks. *Journal of Educational Measurement* 36: 225–33.

Appendix A: ELA and Mathematics Test Configurations

The 2021 Grades 3–8 ELA and Mathematics tests were composed of one session (Session 1) that included MC items. A second session (Session 2), consisting of predominantly CR items, was made available to schools strictly for their optional administration to students and, if administered, student scores for Session 2 test items were to be used for local student assessment purposes only. In addition, the 2021 Tests were reduced in length by the removal of the embedded multiple-choice field test questions. The tables included in Appendix A refer to the test configuration when the operational test forms were previously administered for the 2018 and 2019 Tests, with the embedded field test items removed.

Table A.1. ELA Test Configuration

			Number of Items		
			Multiple-Choice	Constructed-Response	
Grade	Day	Session	Operational	Operational	Total
3	1	1	18	0	18
	2	2	0	7	7
		Total	18	7	25
4	1	1	18	0	18
	2	2	0	7	7
		Total	18	7	25
5	1	1	28	0	28
	2	2	0	7	7
		Total	28	7	35
6	1	1	28	0	28
	2	2	0	7	7
		Total	28	7	35
7	1	1	28	0	28
	2	2	0	8	8
		Total	28	8	36
8	1	1	28	0	28
	2	2	0	8	8
		Total	28	8	36

Table A.2. Mathematics Test Configuration

			Number of Items		
			Multiple-Choice	Constructed-Response	
Grade	Day	Session	Operational	Operational	Total
3	1	1	19	0	19
	2	2	8	7	15
		Total	27	7	34
4	1	1	23	0	23
	2	2	8	7	15
		Total	31	7	38
5	1	1	23	0	23
	2	2	8	7	15
		Total	31	7	38
6	1	1	24	0	24
	2	2	7	8	15

Grade	Day	Session	Number of Items		
			Multiple-Choice	Constructed-Response	
			Operational	Operational	Total
6		Total	31	8	39
7	1	1	26	0	26
	2	2	7	8	15
		Total	33	8	41
8	1	1	26	0	26
	2	2	7	8	15
		Total	33	8	41

Additional details on security, scheduling, classroom organization and preparation, test materials, and administration can be found in the *2021 Teacher's Directions* and the *School Administrator's Manual*, which are accessible online:

- *2021 ELA Teacher's Directions*
  - Grades 3–5: http://www.nysed.gov/common/nysed/files/programs/state-assessment/3-5-ela-paper-teacher-directions-2021.pdf
  - Grades 6–8: http://www.nysed.gov/common/nysed/files/programs/state-assessment/6-8-ela-paper-teacher-directions-2021.pdf
- *2021 Mathematics Teacher's Directions*
  - Grades 3–5: http://www.nysed.gov/common/nysed/files/programs/state-assessment/3-5-math-paper-teacher-directions-2021.pdf
  - Grades 6–8: http://www.nysed.gov/common/nysed/files/programs/state-assessment/6-8-math-paper-teacher-directions-2021.pdf
- *2021 ELA and Mathematics Tests School Administrator's Manual*
  http://www.nysed.gov/common/nysed/files/programs/state-assessment/ei-sam-21ac.pdf
- *2021 ELA and Mathematics Test Guides*
  - https://www.engageny.org/resource/test-guides-english-language-arts-and-mathematics

Appendix B: Operational Item Maps

The following tables show the operational item maps for the 2021 NYSTP Grades 3–8 ELA and Mathematics tests. All the test forms in the Spring 2021 operational tests were previously used in the Spring 2018 and Spring 2019 operational administrations. In addition, only Session 1 of the ELA and Mathematics Tests were required. Session 2 was made available to schools strictly for their optional administration to students and, if administered, student scores for Session 2 test questions were to be used for local student assessment purposes only. Only student scores on the Session 1 test questions contributed to students' total test scores. Additional detail on the standards to which these items align may be found at: http://www.engageny.org/resource/new-york-state-p-12-common-core-learning-standards.

Table B.1. ELA Grade 3 Operational Item Map

Question	Type	Points	Standard	Strand	Subscore
Session 1					
1	Multiple Choice	1	CCSS.ELA-Literacy.L.3.4	Language Standards	Reading
2	Multiple Choice	1	CCSS.ELA-Literacy.RL.3.3	Reading Standards for Literature	Reading
3	Multiple Choice	1	CCSS.ELA-Literacy.RL.3.3	Reading Standards for Literature	Reading
4	Multiple Choice	1	CCSS.ELA-Literacy.RL.3.4	Reading Standards for Literature	Reading
5	Multiple Choice	1	CCSS.ELA-Literacy.RL.3.5	Reading Standards for Literature	Reading
6	Multiple Choice	1	CCSS.ELA-Literacy.RL.3.2	Reading Standards for Literature	Reading
7	Multiple Choice	1	CCSS.ELA-Literacy.RI.3.3	Reading Standards for Informational Text	Reading
8	Multiple Choice	1	CCSS.ELA-Literacy.RI.3.2	Reading Standards for Informational Text	Reading
9	Multiple Choice	1	CCSS.ELA-Literacy.RI.3.7	Reading Standards for Informational Text	Reading
10	Multiple Choice	1	CCSS.ELA-Literacy.RI.3.4	Reading Standards for Informational Text	Reading
11	Multiple Choice	1	CCSS.ELA-Literacy.RI.3.5	Reading Standards for Informational Text	Reading
12	Multiple Choice	1	CCSS.ELA-Literacy.RI.3.3	Reading Standards for Informational Text	Reading
13	Multiple Choice	1	CCSS.ELA-Literacy.RL.3.4	Reading Standards for Literature	Reading
14	Multiple Choice	1	CCSS.ELA-Literacy.RL.3.6	Reading Standards for Literature	Reading
15	Multiple Choice	1	CCSS.ELA-Literacy.RL.3.3	Reading Standards for Literature	Reading

Question	Type	Points	Standard	Strand	Subscore
16	Multiple Choice	1	CCSS.ELA-Literacy.RL.3.2	Reading Standards for Literature	Reading
17	Multiple Choice	1	CCSS.ELA-Literacy.RL.3.5	Reading Standards for Literature	Reading
18	Multiple Choice	1	CCSS.ELA-Literacy.RL.3.2	Reading Standards for Literature	Reading
Session 2					
19	Constructed Response	2	CCSS.ELA-Literacy.RI.3.3	Reading Standards for Informational Text	Writing to Sources
20	Constructed Response	2	CCSS.ELA-Literacy.RI.3.2	Reading Standards for Informational Text	Writing to Sources
21	Constructed Response	2	CCSS.ELA-Literacy.RL.3.5	Reading Standards for Literature	Writing to Sources
22	Constructed Response	2	CCSS.ELA-Literacy.RL.3.3	Reading Standards for Literature	Writing to Sources
23	Constructed Response	2	CCSS.ELA-Literacy.RL.3.2	Reading Standards for Literature	Writing to Sources
24	Constructed Response	2	CCSS.ELA-Literacy.RI.3.4	Reading Standards for Informational Text	Writing to Sources
25	Constructed Response	4	CCSS.ELA-Literacy.RI.3.3	Reading Standards for Informational Text	Writing to Sources

*This item map is intended to identify the primary analytic skills necessary to successfully answer each question. However, each constructed-response question measures proficiencies described in multiple standards, including writing and additional reading and language standards. For example, two-point and four-point constructed-response questions require students to first conduct the analyses described in the mapped standard and then produce written responses that are rated based on writing standards. To gain greater insight into the measurement focus for constructed-response questions, please refer to the rubrics shown in the Educator Guides.

Table B.2. ELA Grade 4 Operational Item Map

Question	Type	Points	Standard	Strand	Subscore
Session 1					
1	Multiple Choice	1	CCSS.ELA-Literacy.RL.4.4	Reading Standards for Literature	Reading
2	Multiple Choice	1	CCSS.ELA-Literacy.L.4.4	Language Standards	Reading
3	Multiple Choice	1	CCSS.ELA-Literacy.RL.4.6	Reading Standards for Literature	Reading
4	Multiple Choice	1	CCSS.ELA-Literacy.RL.4.3	Reading Standards for Literature	Reading
5	Multiple Choice	1	CCSS.ELA-Literacy.RL.4.3	Reading Standards for Literature	Reading
6	Multiple Choice	1	CCSS.ELA-Literacy.RL.4.2	Reading Standards for Literature	Reading
7	Multiple Choice	1	CCSS.ELA-Literacy.RI.4.4	Reading Standards for Informational Text	Reading

Question	Type	Points	Standard	Strand	Subscore
8	Multiple Choice	1	CCSS.ELA-Literacy.RI.4.3	Reading Standards for Informational Text	Reading
9	Multiple Choice	1	CCSS.ELA-Literacy.RI.4.3	Reading Standards for Informational Text	Reading
10	Multiple Choice	1	CCSS.ELA-Literacy.RI.4.2	Reading Standards for Informational Text	Reading
11	Multiple Choice	1	CCSS.ELA-Literacy.RI.4.5	Reading Standards for Informational Text	Reading
12	Multiple Choice	1	CCSS.ELA-Literacy.RI.4.7	Reading Standards for Informational Text	Reading
13	Multiple Choice	1	CCSS.ELA-Literacy.RL.4.2	Reading Standards for Literature	Reading
14	Multiple Choice	1	CCSS.ELA-Literacy.RL.4.4	Reading Standards for Literature	Reading
15	Multiple Choice	1	CCSS.ELA-Literacy.L.4.4	Language Standards	Reading
16	Multiple Choice	1	CCSS.ELA-Literacy.RL.4.3	Reading Standards for Literature	Reading
17	Multiple Choice	1	CCSS.ELA-Literacy.RL.4.3	Reading Standards for Literature	Reading
18	Multiple Choice	1	CCSS.ELA-Literacy.RL.4.2	Reading Standards for Literature	Reading
Session 2					
19	Constructed Response	2	CCSS.ELA-Literacy.RL.4.4	Reading Standards for Literature	Writing to Sources
20	Constructed Response	2	CCSS.ELA-Literacy.RL.4.2	Reading Standards for Literature	Writing to Sources
21	Constructed Response	2	CCSS.ELA-Literacy.RL.4.6	Reading Standards for Literature	Writing to Sources
22	Constructed Response	2	CCSS.ELA-Literacy.RI.4.3	Reading Standards for Informational Text	Writing to Sources
23	Constructed Response	2	CCSS.ELA-Literacy.RI.4.2	Reading Standards for Informational Text	Writing to Sources
24	Constructed Response	2	CCSS.ELA-Literacy.RI.4.3	Reading Standards for Informational Text	Writing to Sources
25	Constructed Response	4	CCSS.ELA-Literacy.RI.4.3	Reading Standards for Informational Text	Writing to Sources

*This item map is intended to identify the primary analytic skills necessary to successfully answer each question. However, each constructed-response question measures proficiencies described in multiple standards, including writing and additional reading and language standards. For example, two point and four point constructed-response questions require students to first conduct the analyses described in the mapped standard and then produce written responses that are rated based on writing standards. To gain greater insight into the measurement focus for constructed-response questions please refer to the rubrics shown in the Educator Guides.

Table B.3. ELA Grade 5 Operational Item Map

Question	Type	Points	Standard	Strand	Subscore
Session 1					
1	Multiple Choice	1	CCSS.ELA-Literacy.RL.5.4	Reading Standards for Literature	Reading
2	Multiple Choice	1	CCSS.ELA-Literacy.RL.5.5	Reading Standards for Literature	Reading
3	Multiple Choice	1	CCSS.ELA-Literacy.RL.5.2	Reading Standards for Literature	Reading
4	Multiple Choice	1	CCSS.ELA-Literacy.RL.5.6	Reading Standards for Literature	Reading
5	Multiple Choice	1	CCSS.ELA-Literacy.RL.5.4	Reading Standards for Literature	Reading
6	Multiple Choice	1	CCSS.ELA-Literacy.RL.5.3	Reading Standards for Literature	Reading
7	Multiple Choice	1	CCSS.ELA-Literacy.RL.5.3	Reading Standards for Literature	Reading
8	Multiple Choice	1	CCSS.ELA-Literacy.RI.5.4	Reading Standards for Informational Text	Reading
9	Multiple Choice	1	CCSS.ELA-Literacy.L.5.4	Language Standards	Reading
10	Multiple Choice	1	CCSS.ELA-Literacy.RI.5.2	Reading Standards for Informational Text	Reading
11	Multiple Choice	1	CCSS.ELA-Literacy.RI.5.6	Reading Standards for Informational Text	Reading
12	Multiple Choice	1	CCSS.ELA-Literacy.RI.5.3	Reading Standards for Informational Text	Reading
13	Multiple Choice	1	CCSS.ELA-Literacy.RI.5.3	Reading Standards for Informational Text	Reading
14	Multiple Choice	1	CCSS.ELA-Literacy.RI.5.3	Reading Standards for Informational Text	Reading
15	Multiple Choice	1	CCSS.ELA-Literacy.RL.5.5	Reading Standards for Literature	Reading
16	Multiple Choice	1	CCSS.ELA-Literacy.RL.5.4	Reading Standards for Literature	Reading
17	Multiple Choice	1	CCSS.ELA-Literacy.RL.5.3	Reading Standards for Literature	Reading
18	Multiple Choice	1	CCSS.ELA-Literacy.RL.5.3	Reading Standards for Literature	Reading
19	Multiple Choice	1	CCSS.ELA-Literacy.RL.5.3	Reading Standards for Literature	Reading
20	Multiple Choice	1	CCSS.ELA-Literacy.RL.5.2	Reading Standards for Literature	Reading
21	Multiple Choice	1	CCSS.ELA-Literacy.RL.5.2	Reading Standards for Literature	Reading
22	Multiple Choice	1	CCSS.ELA-Literacy.RI.5.3	Reading Standards for Informational Text	Reading

Question	Type	Points	Standard	Strand	Subscore
23	Multiple Choice	1	CCSS.ELA-Literacy.L.5.5	Language Standards	Reading
24	Multiple Choice	1	CCSS.ELA-Literacy.RI.5.3	Reading Standards for Informational Text	Reading
25	Multiple Choice	1	CCSS.ELA-Literacy.RI.5.3	Reading Standards for Informational Text	Reading
26	Multiple Choice	1	CCSS.ELA-Literacy.RI.5.6	Reading Standards for Informational Text	Reading
27	Multiple Choice	1	CCSS.ELA-Literacy.RI.5.2	Reading Standards for Informational Text	Reading
28	Multiple Choice	1	CCSS.ELA-Literacy.RI.5.3	Reading Standards for Informational Text	Reading
Session 2					
29	Constructed Response	2	CCSS.ELA-Literacy.RI.5.2	Reading Standards for Informational Text	Writing to Sources
30	Constructed Response	2	CCSS.ELA-Literacy.RI.5.3	Reading Standards for Informational Text	Writing to Sources
31	Constructed Response	2	CCSS.ELA-Literacy.RI.5.3	Reading Standards for Informational Text	Writing to Sources
32	Constructed Response	2	CCSS.ELA-Literacy.RL.5.3	Reading Standards for Literature	Writing to Sources
33	Constructed Response	2	CCSS.ELA-Literacy.RL.5.4	Reading Standards for Literature	Writing to Sources
34	Constructed Response	2	CCSS.ELA-Literacy.RL.5.5	Reading Standards for Literature	Writing to Sources
35	Constructed Response	4	CCSS.ELA-Literacy.RL.5.6	Reading Standards for Literature	Writing to Sources

*This item map is intended to identify the primary analytic skills necessary to successfully answer each question. However, each constructed-response question measures proficiencies described in multiple standards, including writing and additional reading and language standards. For example, two-point and four-point constructed-response questions require students to first conduct the analyses described in the mapped standard and then produce written responses that are rated based on writing standards. To gain greater insight into the measurement focus for constructed-response questions, please refer to the rubrics shown in the Educator Guides.

Table B.4. ELA Grade 6 Operational Item Map

Question	Type	Points	Standard	Strand	Subscore
Session 1					
1	Multiple Choice	1	CCSS.ELA-Literacy.RL.6.3	Reading Standards for Literature	Reading
2	Multiple Choice	1	CCSS.ELA-Literacy.RL.6.5	Reading Standards for Literature	Reading
3	Multiple Choice	1	CCSS.ELA-Literacy.RL.6.4	Reading Standards for Literature	Reading
4	Multiple Choice	1	CCSS.ELA-Literacy.RL.6.3	Reading Standards for Literature	Reading

Question	Type	Points	Standard	Strand	Subscore
5	Multiple Choice	1	CCSS.ELA-Literacy.RL.6.6	Reading Standards for Literature	Reading
6	Multiple Choice	1	CCSS.ELA-Literacy.RL.6.2	Reading Standards for Literature	Reading
7	Multiple Choice	1	CCSS.ELA-Literacy.RL.6.2	Reading Standards for Literature	Reading
8	Multiple Choice	1	CCSS.ELA-Literacy.RI.6.3	Reading Standards for Informational Text	Reading
9	Multiple Choice	1	CCSS.ELA-Literacy.RI.6.8	Reading Standards for Informational Text	Reading
10	Multiple Choice	1	CCSS.ELA-Literacy.L.6.4	Language Standards	Reading
11	Multiple Choice	1	CCSS.ELA-Literacy.RI.6.2	Reading Standards for Informational Text	Reading
12	Multiple Choice	1	CCSS.ELA-Literacy.RI.6.7	Reading Standards for Informational Text	Reading
13	Multiple Choice	1	CCSS.ELA-Literacy.RI.6.5	Reading Standards for Informational Text	Reading
14	Multiple Choice	1	CCSS.ELA-Literacy.RI.6.2	Reading Standards for Informational Text	Reading
15	Multiple Choice	1	CCSS.ELA-Literacy.RL.6.3	Reading Standards for Literature	Reading
16	Multiple Choice	1	CCSS.ELA-Literacy.RL.6.2	Reading Standards for Literature	Reading
17	Multiple Choice	1	CCSS.ELA-Literacy.RL.6.3	Reading Standards for Literature	Reading
18	Multiple Choice	1	CCSS.ELA-Literacy.RL.6.4	Reading Standards for Literature	Reading
19	Multiple Choice	1	CCSS.ELA-Literacy.RL.6.3	Reading Standards for Literature	Reading
20	Multiple Choice	1	CCSS.ELA-Literacy.RL.6.2	Reading Standards for Literature	Reading
21	Multiple Choice	1	CCSS.ELA-Literacy.RL.6.6	Reading Standards for Literature	Reading
22	Multiple Choice	1	CCSS.ELA-Literacy.RI.6.3	Reading Standards for Informational Text	Reading
23	Multiple Choice	1	CCSS.ELA-Literacy.RI.6.4	Reading Standards for Informational Text	Reading
24	Multiple Choice	1	CCSS.ELA-Literacy.RI.6.7	Reading Standards for Informational Text	Reading
25	Multiple Choice	1	CCSS.ELA-Literacy.RI.6.5	Reading Standards for Informational Text	Reading
26	Multiple Choice	1	CCSS.ELA-Literacy.RI.6.2	Reading Standards for Informational Text	Reading
27	Multiple Choice	1	CCSS.ELA-Literacy.RI.6.6	Reading Standards for Informational Text	Reading
28	Multiple Choice	1	CCSS.ELA-Literacy.RI.6.2	Reading Standards for Informational Text	Reading

Question	Type	Points	Standard	Strand	Subscore
Session 2					
29	Constructed Response	2	CCSS.ELA-Literacy.RL.6.4	Reading Standards for Literature	Writing to Sources
30	Constructed Response	2	CCSS.ELA-Literacy.RL.6.6	Reading Standards for Literature	Writing to Sources
31	Constructed Response	2	CCSS.ELA-Literacy.RL.6.5	Reading Standards for Literature	Writing to Sources
32	Constructed Response	2	CCSS.ELA-Literacy.RI.6.2	Reading Standards for Informational Text	Writing to Sources
33	Constructed Response	2	CCSS.ELA-Literacy.RI.6.3	Reading Standards for Informational Text	Writing to Sources
34	Constructed Response	2	CCSS.ELA-Literacy.RI.6.2	Reading Standards for Informational Text	Writing to Sources
35	Constructed Response	4	CCSS.ELA-Literacy.RI.6.6	Reading Standards for Informational Text	Writing to Sources

*This item map is intended to identify the primary analytic skills necessary to successfully answer each question. However, each constructed-response question measures proficiencies described in multiple standards, including writing and additional reading and language standards. For example, two point and four point constructed-response questions require students to first conduct the analyses described in the mapped standard and then produce written responses that are rated based on writing standards. To gain greater insight into the measurement focus for constructed-response questions please refer to the rubrics shown in the Educator Guides.

Table B.5. ELA Grade 7 Operational Item Map

Question	Type	Points	Standard	Strand	Subscore
Session 1					
1	Multiple Choice	1	CCSS.ELA-Literacy.L.7.4	Language Standards	Reading
2	Multiple Choice	1	CCSS.ELA-Literacy.RL.7.4	Reading Standards for Literature	Reading
3	Multiple Choice	1	CCSS.ELA-Literacy.RL.7.3	Reading Standards for Literature	Reading
4	Multiple Choice	1	CCSS.ELA-Literacy.RL.7.3	Reading Standards for Literature	Reading
5	Multiple Choice	1	CCSS.ELA-Literacy.RL.7.3	Reading Standards for Literature	Reading
6	Multiple Choice	1	CCSS.ELA-Literacy.RL.7.4	Reading Standards for Literature	Reading
7	Multiple Choice	1	CCSS.ELA-Literacy.RL.7.2	Reading Standards for Literature	Reading
8	Multiple Choice	1	CCSS.ELA-Literacy.RI.7.3	Reading Standards for Informational Text	Reading
9	Multiple Choice	1	CCSS.ELA-Literacy.RI.7.4	Reading Standards for Informational Text	Reading
10	Multiple Choice	1	CCSS.ELA-Literacy.RI.7.5	Reading Standards for Informational Text	Reading

Question	Type	Points	Standard	Strand	Subscore
11	Multiple Choice	1	CCSS.ELA-Literacy.RI.7.2	Reading Standards for Informational Text	Reading
12	Multiple Choice	1	CCSS.ELA-Literacy.RI.7.8	Reading Standards for Informational Text	Reading
13	Multiple Choice	1	CCSS.ELA-Literacy.RI.7.6	Reading Standards for Informational Text	Reading
14	Multiple Choice	1	CCSS.ELA-Literacy.RI.7.2	Reading Standards for Informational Text	Reading
15	Multiple Choice	1	CCSS.ELA-Literacy.RL.7.4	Reading Standards for Literature	Reading
16	Multiple Choice	1	CCSS.ELA-Literacy.RL.7.3	Reading Standards for Literature	Reading
17	Multiple Choice	1	CCSS.ELA-Literacy.RL.7.2	Reading Standards for Literature	Reading
18	Multiple Choice	1	CCSS.ELA-Literacy.RL.7.2	Reading Standards for Literature	Reading
19	Multiple Choice	1	CCSS.ELA-Literacy.RL.7.2	Reading Standards for Literature	Reading
20	Multiple Choice	1	CCSS.ELA-Literacy.RL.7.3	Reading Standards for Literature	Reading
21	Multiple Choice	1	CCSS.ELA-Literacy.RL.7.6	Reading Standards for Literature	Reading
22	Multiple Choice	1	CCSS.ELA-Literacy.RI.7.2	Reading Standards for Informational Text	Reading
23	Multiple Choice	1	CCSS.ELA-Literacy.RI.7.4	Reading Standards for Informational Text	Reading
24	Multiple Choice	1	CCSS.ELA-Literacy.RI.7.4	Reading Standards for Informational Text	Reading
25	Multiple Choice	1	CCSS.ELA-Literacy.RI.7.3	Reading Standards for Informational Text	Reading
26	Multiple Choice	1	CCSS.ELA-Literacy.RI.7.8	Reading Standards for Informational Text	Reading
27	Multiple Choice	1	CCSS.ELA-Literacy.RI.7.3	Reading Standards for Informational Text	Reading
28	Multiple Choice	1	CCSS.ELA-Literacy.RI.7.3	Reading Standards for Informational Text	Reading
Session 2					
29	Constructed Response	2	CCSS.ELA-Literacy.RI.7.6	Reading Standards for Informational Text	Writing to Sources
30	Constructed Response	2	CCSS.ELA-Literacy.RI.7.5	Reading Standards for Informational Text	Writing to Sources
31	Constructed Response	2	CCSS.ELA-Literacy.RI.7.5	Reading Standards for Informational Text	Writing to Sources
32	Constructed Response	2	CCSS.ELA-Literacy.RL.7.3	Reading Standards for Literature	Writing to Sources
33	Constructed Response	2	CCSS.ELA-Literacy.RL.7.2	Reading Standards for Literature	Writing to Sources

Question	Type	Points	Standard	Strand	Subscore
34	Constructed Response	2	CCSS.ELA-Literacy.RL.7.4	Reading Standards for Literature	Writing to Sources
35	Constructed Response	2	CCSS.ELA-Literacy.RL.7.6	Reading Standards for Literature	Writing to Sources
36	Constructed Response	4	CCSS.ELA-Literacy.RL.7.3	Reading Standards for Literature	Writing to Sources

*This item map is intended to identify the primary analytic skills necessary to successfully answer each question. However, each constructed-response question measures proficiencies described in multiple standards, including writing and additional reading and language standards. For example, two point and four point constructed-response questions require students to first conduct the analyses described in the mapped standard and then produce written responses that are rated based on writing standards. To gain greater insight into the measurement focus for constructed-response questions please refer to the rubrics shown in the Educator Guides.

Table B.6. ELA Grade 8 Operational Item Map

Question	Type	Points	Standard	Strand	Subscore
Session 1					
1	Multiple Choice	1	CCSS.ELA-Literacy.RL.8.3	Reading Standards for Literature	Reading
2	Multiple Choice	1	CCSS.ELA-Literacy.RL.8.4	Reading Standards for Literature	Reading
3	Multiple Choice	1	CCSS.ELA-Literacy.RL.8.3	Reading Standards for Literature	Reading
4	Multiple Choice	1	CCSS.ELA-Literacy.RL.8.3	Reading Standards for Literature	Reading
5	Multiple Choice	1	CCSS.ELA-Literacy.L.8.4	Language Standards	Reading
6	Multiple Choice	1	CCSS.ELA-Literacy.RL.8.6	Reading Standards for Literature	Reading
7	Multiple Choice	1	CCSS.ELA-Literacy.RL.8.2	Reading Standards for Literature	Reading
8	Multiple Choice	1	CCSS.ELA-Literacy.RI.8.4	Reading Standards for Informational Text	Reading
9	Multiple Choice	1	CCSS.ELA-Literacy.RI.8.8	Reading Standards for Informational Text	Reading
10	Multiple Choice	1	CCSS.ELA-Literacy.RI.8.5	Reading Standards for Informational Text	Reading
11	Multiple Choice	1	CCSS.ELA-Literacy.L.8.4	Language Standards	Reading
12	Multiple Choice	1	CCSS.ELA-Literacy.RI.8.3	Reading Standards for Informational Text	Reading
13	Multiple Choice	1	CCSS.ELA-Literacy.RI.8.2	Reading Standards for Informational Text	Reading
14	Multiple Choice	1	CCSS.ELA-Literacy.RI.8.6	Reading Standards for Informational Text	Reading
15	Multiple Choice	1	CCSS.ELA-Literacy.RL.8.4	Reading Standards for Literature	Reading

Question	Type	Points	Standard	Strand	Subscore
16	Multiple Choice	1	CCSS.ELA-Literacy.RL.8.2	Reading Standards for Literature	Reading
17	Multiple Choice	1	CCSS.ELA-Literacy.RL.8.3	Reading Standards for Literature	Reading
18	Multiple Choice	1	CCSS.ELA-Literacy.RL.8.3	Reading Standards for Literature	Reading
19	Multiple Choice	1	CCSS.ELA-Literacy.RL.8.4	Reading Standards for Literature	Reading
20	Multiple Choice	1	CCSS.ELA-Literacy.RL.8.2	Reading Standards for Literature	Reading
21	Multiple Choice	1	CCSS.ELA-Literacy.RL.8.6	Reading Standards for Literature	Reading
22	Multiple Choice	1	CCSS.ELA-Literacy.RI.8.4	Reading Standards for Informational Text	Reading
23	Multiple Choice	1	CCSS.ELA-Literacy.RI.8.4	Reading Standards for Informational Text	Reading
24	Multiple Choice	1	CCSS.ELA-Literacy.RI.8.5	Reading Standards for Informational Text	Reading
25	Multiple Choice	1	CCSS.ELA-Literacy.RI.8.6	Reading Standards for Informational Text	Reading
26	Multiple Choice	1	CCSS.ELA-Literacy.RI.8.3	Reading Standards for Informational Text	Reading
27	Multiple Choice	1	CCSS.ELA-Literacy.RI.8.3	Reading Standards for Informational Text	Reading
28	Multiple Choice	1	CCSS.ELA-Literacy.RI.8.2	Reading Standards for Informational Text	Reading
Session 2					
29	Constructed Response	2	CCSS.ELA-Literacy.RL.8.4	Reading Standards for Literature	Writing to Sources
30	Constructed Response	2	CCSS.ELA-Literacy.RL.8.6	Reading Standards for Literature	Writing to Sources
31	Constructed Response	2	CCSS.ELA-Literacy.RL.8.2	Reading Standards for Literature	Writing to Sources
32	Constructed Response	2	CCSS.ELA-Literacy.RI.8.4	Reading Standards for Informational Text	Writing to Sources
33	Constructed Response	2	CCSS.ELA-Literacy.RI.8.5	Reading Standards for Informational Text	Writing to Sources
34	Constructed Response	2	CCSS.ELA-Literacy.RI.8.2	Reading Standards for Informational Text	Writing to Sources
35	Constructed Response	2	CCSS.ELA-Literacy.RI.8.3	Reading Standards for Informational Text	Writing to Sources
36	Constructed Response	4	CCSS.ELA-Literacy.RI.8.8	Reading Standards for Informational Text	Writing to Sources

*This item map is intended to identify the primary analytic skills necessary to successfully answer each question. However, each constructed-response question measures proficiencies described in multiple standards, including writing and additional reading and language standards. For example, two-point and four-point constructed-response questions require students to first conduct the analyses described in the mapped standard and then produce written responses that are rated based on writing standards. To gain greater insight into the measurement focus for constructed-response questions, please refer to the rubrics shown in the Educator Guides.

Table B.7. Mathematics Grade 3 Operational Item Map

Question	Type	Points	Standard	Strand	Subscore
Session 1					
1	Multiple Choice	1	CCSS.Math.Content.3.OA.B.5	Operations and Algebraic Thinking	Operations and Algebraic Thinking
2	Multiple Choice	1	CCSS.Math.Content.3.NBT.A.1	Numbers and Operations in Base Ten	
3	Multiple Choice	1	CCSS.Math.Content.3.OA.A.4	Operations and Algebraic Thinking	Operations and Algebraic Thinking
4	Multiple Choice	1	CCSS.Math.Content.3.NF.A.3b	Number and Operations— Fractions	Number and Operations—Fractions
5	Multiple Choice	1	CCSS.Math.Content.3.OA.A.1	Operations and Algebraic Thinking	Operations and Algebraic Thinking
6	Multiple Choice	1	CCSS.Math.Content.3.MD.C.5b	Measurement and Data	Measurement and Data
7	Multiple Choice	1	CCSS.Math.Content.3.OA.D.8	Operations and Algebraic Thinking	Operations and Algebraic Thinking
8	Multiple Choice	1	CCSS.Math.Content.3.NBT.A.3	Numbers and Operations in Base Ten	
9	Multiple Choice	1	CCSS.Math.Content.3.OA.B.6	Operations and Algebraic Thinking	Operations and Algebraic Thinking
10	Multiple Choice	1	CCSS.Math.Content.3.OA.A.2	Operations and Algebraic Thinking	Operations and Algebraic Thinking
11	Multiple Choice	1	CCSS.Math.Content.3.G.A.2	Geometry	
12	Multiple Choice	1	CCSS.Math.Content.3.MD.A.1	Measurement and Data	Measurement and Data
13	Multiple Choice	1	CCSS.Math.Content.3.NF.A.2b	Number and Operations— Fractions	Number and Operations—Fractions
14	Multiple Choice	1	CCSS.Math.Content.3.OA.A.3	Operations and Algebraic Thinking	Operations and Algebraic Thinking
15	Multiple Choice	1	CCSS.Math.Content.3.OA.D.9	Operations and Algebraic Thinking	Operations and Algebraic Thinking
16	Multiple Choice	1	CCSS.Math.Content.3.NF.A.2a	Number and Operations— Fractions	Number and Operations—Fractions
17	Multiple Choice		CCSS.Math.Content.3.MD.A.2	Measurement and Data	Measurement and Data
18	Multiple Choice	1	CCSS.Math.Content.3.OA.B.5	Operations and Algebraic Thinking	Operations and Algebraic Thinking
19	Multiple Choice	1	CCSS.Math.Content.3.NF.A.3d	Number and Operations— Fractions	Number and Operations—Fractions

Question	Type	Points	Standard	Strand	Subscore
Session 2					
20	Multiple Choice	1	CCSS.Math.Content.3.OA.A.4	Operations and Algebraic Thinking	Operations and Algebraic Thinking
21	Multiple Choice	1	CCSS.Math.Content.3.NBT.A.3	Numbers and Operations in Base Ten	
22	Multiple Choice	1	CCSS.Math.Content.3.NF.A.3a	Number and Operations— Fractions	Number and Operations—Fractions
23	Multiple Choice	1	CCSS.Math.Content.3.OA.A.2	Operations and Algebraic Thinking	Operations and Algebraic Thinking
24	Multiple Choice	1	CCSS.Math.Content.3.OA.D.9	Operations and Algebraic Thinking	Operations and Algebraic Thinking
25	Multiple Choice	1	CCSS.Math.Content.3.MD.C.7d	Measurement and Data	Measurement and Data
26	Multiple Choice	1	CCSS.Math.Content.3.G.A.2	Geometry	
27	Multiple Choice	1	CCSS.Math.Content.3.NF.A.3c	Number and Operations— Fractions	Number and Operations—Fractions
28	Constructed Response	2	CCSS.Math.Content.3.MD.A.1	Measurement and Data	Measurement and Data
29	Constructed Response	2	CCSS.Math.Content.3.OA.A.1	Operations and Algebraic Thinking	Operations and Algebraic Thinking
30	Constructed Response	2	CCSS.Math.Content.3.NF.A.1	Number and Operations— Fractions	Number and Operations—Fractions
31	Constructed Response	2	CCSS.Math.Content.3.MD.B.3	Measurement and Data	Measurement and Data
32	Constructed Response	2	CCSS.Math.Content.3.OA.A.3	Operations and Algebraic Thinking	Operations and Algebraic Thinking
33	Constructed Response	2	CCSS.Math.Content.3.MD.C.7b	Measurement and Data	Measurement and Data
34	Constructed Response	3	CCSS.Math.Content.3.OA.D.8	Operations and Algebraic Thinking	Operations and Algebraic Thinking

*This item map is intended to identify the primary analytic skills necessary to successfully answer each question. However, some questions measure proficiencies described in multiple standards, including a balanced combination of procedural and conceptual understanding.

Table B.8. Mathematics Grade 4 Operational Item Map

Question	Type	Points	Standard	Cluster	Subscore
Session 1					
1	Multiple Choice	1	CCSS.Math.Content.4.NF.B.4c	Number and Operations - Fractions	Number and Operations - Fractions

Question	Type	Points	Standard	Cluster	Subscore
2	Multiple Choice	1	CCSS.Math.Content.4.OA.B.4	Operations and Algebraic Thinking	Operations and Algebraic Thinking
3	Multiple Choice	1	CCSS.Math.Content.4.NF.A.2	Number and Operations - Fractions	Number and Operations - Fractions
4	Multiple Choice	1	CCSS.Math.Content.4.OA.A.2	Operations and Algebraic Thinking	Operations and Algebraic Thinking
5	Multiple Choice	1	CCSS.Math.Content.4.G.A.1	Geometry	
6	Multiple Choice	1	CCSS.Math.Content.4.OA.A.3	Operations and Algebraic Thinking	Operations and Algebraic Thinking
7	Multiple Choice	1	CCSS.Math.Content.4.NF.A.1	Number and Operations - Fractions	Number and Operations - Fractions
8	Multiple Choice	1	CCSS.Math.Content.4.MD.C.7	Measurement and Data	
9	Multiple Choice	1	CCSS.Math.Content.4.NF.A.2	Number and Operations - Fractions	Number and Operations - Fractions
10	Multiple Choice	1	CCSS.Math.Content.4.NF.B.3c	Number and Operations - Fractions	Number and Operations - Fractions
11	Multiple Choice	1	CCSS.Math.Content.4.NBT.B.5	Number and Operations in Base Ten	Number and Operations in Base Ten
12	Multiple Choice	1	CCSS.Math.Content.4.NBT.B.6	Number and Operations in Base Ten	Number and Operations in Base Ten
13	Multiple Choice	1	CCSS.Math.Content.4.OA.A.3	Operations and Algebraic Thinking	Operations and Algebraic Thinking
14	Multiple Choice	1	CCSS.Math.Content.4.OA.A.1	Operations and Algebraic Thinking	Operations and Algebraic Thinking
15	Multiple Choice	1	CCSS.Math.Content.4.NBT.A.3	Number and Operations in Base Ten	Number and Operations in Base Ten
16	Multiple Choice	1	CCSS.Math.Content.4.MD.A.3	Measurement and Data	
17	Multiple Choice	1	CCSS.Math.Content.4.MD.B.4	Measurement and Data	
18	Multiple Choice	1	CCSS.Math.Content.4.NBT.A.1	Number and Operations in Base Ten	Number and Operations in Base Ten
19	Multiple Choice	1	CCSS.Math.Content.4.OA.C.5	Operations and Algebraic Thinking	Operations and Algebraic Thinking
20	Multiple Choice	1	CCSS.Math.Content.4.NF.B.3a	Number and Operations - Fractions	Number and Operations - Fractions
21	Multiple Choice	1	CCSS.Math.Content.4.G.A.3	Geometry	
22	Multiple Choice	1	CCSS.Math.Content.4.MD.C.5a	Measurement and Data	
23	Multiple Choice	1	CCSS.Math.Content.4.OA.A.1	Operations and Algebraic Thinking	Operations and Algebraic Thinking
Session 2					
24	Multiple Choice	1	CCSS.Math.Content.4.NF.B.4a	Number and Operations - Fractions	Number and Operations - Fractions

Question	Type	Points	Standard	Cluster	Subscore
25	Multiple Choice	1	CCSS.Math.Content.4.OA.A.2	Operations and Algebraic Thinking	Operations and Algebraic Thinking
26	Multiple Choice	1	CCSS.Math.Content.4.NBT.B.6	Number and Operations in Base Ten	Number and Operations in Base Ten
27	Multiple Choice	1	CCSS.Math.Content.4.MD.C.5b	Measurement and Data	
28	Multiple Choice	1	CCSS.Math.Content.4.OA.A.1	Operations and Algebraic Thinking	Operations and Algebraic Thinking
29	Multiple Choice	1	CCSS.Math.Content.4.MD.C.6	Measurement and Data	
30	Multiple Choice	1	CCSS.Math.Content.4.NF.A.1	Number and Operations - Fractions	Number and Operations - Fractions
31	Multiple Choice	1	CCSS.Math.Content.3.MD.B.4	Measurement and Data	
32	Constructed Response	2	CCSS.Math.Content.4.NF.B.4b	Number and Operations - Fractions	Number and Operations - Fractions
33	Constructed Response	2	CCSS.Math.Content.4.NBT.A.2	Number and Operations in Base Ten	Number and Operations in Base Ten
34	Constructed Response	2	CCSS.Math.Content.4.G.A.2	Geometry	
35	Constructed Response	2	CCSS.Math.Content.4.NF.B.3d	Number and Operations - Fractions	Number and Operations - Fractions
36	Constructed Response	2	CCSS.Math.Content.4.NBT.A.1	Number and Operations in Base Ten	Number and Operations in Base Ten
37	Constructed Response	2	CCSS.Math.Content.4.MD.A.3	Measurement and Data	
38	Constructed Response	3	CCSS.Math.Content.4.NBT.B.5	Number and Operations in Base Ten	Number and Operations in Base Ten

*This item map is intended to identify the primary analytic skills necessary to successfully answer each question. However, some questions measure proficiencies described in multiple standards, including a balanced combination of procedural and conceptual understanding.

Table B.9. Mathematics Grade 5 Operational Item Map

Question	Type	Points	Standard	Cluster	Subscore
Session 1					
1	Multiple Choice	1	CCSS.Math.Content.5.MD.C.5b	Measurement and Data	Measurement and Data
2	Multiple Choice	1	CCSS.Math.Content.4.NF.C.5	Number and Operations - Fractions	Number and Operations - Fractions
3	Multiple Choice	1	CCSS.Math.Content.5.NF.B.6	Number and Operations - Fractions	Number and Operations - Fractions
4	Multiple Choice	1	CCSS.Math.Content.4.NF.C.6	Number and Operations in Base Ten	Number and Operations in Base Ten
5	Multiple Choice	1	CCSS.Math.Content.5.NBT.A.3b	Number and Operations in Base Ten	Number and Operations in Base Ten

Question	Type	Points	Standard	Cluster	Subscore
6	Multiple Choice	1	CCSS.Math.Content.5.OA.A.2	Operations and Algebraic Thinking	
7	Multiple Choice	1	CCSS.Math.Content.5.MD.C.4	Measurement and Data	Measurement and Data
8	Multiple Choice	1	CCSS.Math.Content.5.OA.A.1	Operations and Algebraic Thinking	
9	Multiple Choice	1	CCSS.Math.Content.5.MD.A.1	Measurement and Data	Measurement and Data
10	Multiple Choice	1	CCSS.Math.Content.5.NF.B.4b	Number and Operations - Fractions	Number and Operations - Fractions
11	Multiple Choice	1	CCSS.Math.Content.4.MD.A.2	Measurement and Data	Measurement and Data
12	Multiple Choice	1	CCSS.Math.Content.5.NBT.B.6	Number and Operations in Base Ten	Number and Operations in Base Ten
13	Multiple Choice	1	CCSS.Math.Content.5.MD.C.5a	Measurement and Data	Measurement and Data
14	Multiple Choice	1	CCSS.Math.Content.5.NBT.A.4	Number and Operations in Base Ten	Number and Operations in Base Ten
15	Multiple Choice	1	CCSS.Math.Content.5.NF.B.7c	Number and Operations - Fractions	Number and Operations - Fractions
16	Multiple Choice	1	CCSS.Math.Content.5.NF.A.2	Number and Operations - Fractions	Number and Operations - Fractions
17	Multiple Choice	1	CCSS.Math.Content.5.NF.B.6	Number and Operations - Fractions	Number and Operations - Fractions
18	Multiple Choice	1	CCSS.Math.Content.5.NF.B.5a	Number and Operations - Fractions	Number and Operations - Fractions
19	Multiple Choice	1	CCSS.Math.Content.5.NF.B.7a	Number and Operations - Fractions	Number and Operations - Fractions
20	Multiple Choice	1	CCSS.Math.Content.5.NBT.B.7	Number and Operations in Base Ten	Number and Operations in Base Ten
21	Multiple Choice	1	CCSS.Math.Content.5.NBT.B.6	Number and Operations in Base Ten	Number and Operations in Base Ten
22	Multiple Choice	1	CCSS.Math.Content.5.MD.B.2	Measurement and Data	Measurement and Data
23	Multiple Choice	1	CCSS.Math.Content.5.MD.C.5c	Measurement and Data	Measurement and Data
Session 2					
24	Multiple Choice	1	CCSS.Math.Content.5.G.B.3	Geometry	
25	Multiple Choice	1	CCSS.Math.Content.5.NF.A.1	Number and Operations - Fractions	Number and Operations - Fractions
26	Multiple Choice	1	CCSS.Math.Content.4.MD.A.1	Measurement and Data	Measurement and Data
27	Multiple Choice	1	CCSS.Math.Content.5.NF.A.1	Number and Operations - Fractions	Number and Operations - Fractions
28	Multiple Choice	1	CCSS.Math.Content.5.NBT.B.7	Number and Operations in Base Ten	Number and Operations in Base Ten

Question	Type	Points	Standard	Cluster	Subscore
29	Multiple Choice	1	CCSS.Math.Content.5.NBT.A.3a	Number and Operations in Base Ten	Number and Operations in Base Ten
30	Multiple Choice	1	CCSS.Math.Content.5.NF.B.6	Number and Operations - Fractions	Number and Operations - Fractions
31	Multiple Choice	1	CCSS.Math.Content.5.NF.B.4a	Number and Operations - Fractions	Number and Operations - Fractions
32	Constructed Response	2	CCSS.Math.Content.5.MD.C.5c	Measurement and Data	Measurement and Data
33	Constructed Response	2	CCSS.Math.Content.5.NF.A.2	Number and Operations - Fractions	Number and Operations - Fractions
34	Constructed Response	2	CCSS.Math.Content.5.NBT.A.1	Number and Operations in Base Ten	Number and Operations in Base Ten
35	Constructed Response	2	CCSS.Math.Content.5.NF.B.7c	Number and Operations - Fractions	Number and Operations - Fractions
36	Constructed Response	2	CCSS.Math.Content.5.MD.A.1	Measurement and Data	Measurement and Data
37	Constructed Response	2	CCSS.Math.Content.5.NF.B.5b	Number and Operations - Fractions	Number and Operations - Fractions
38	Constructed Response	3	CCSS.Math.Content.5.NBT.B.7	Number and Operations in Base Ten	Number and Operations in Base Ten

*This item map is intended to identify the primary analytic skills necessary to successfully answer each question. However, some questions measure proficiencies described in multiple standards, including a balanced combination of procedural and conceptual understanding.

Table B.10. Mathematics Grade 6 Operational Item Map

Question	Type	Points	Standard	Strand	Subscore
Session 1					
1	Multiple Choice	1	CCSS.Math.Content.6.EE.B.5	Expressions and Equations	Expressions and Equations
2	Multiple Choice	1	CCSS.Math.Content.6.RP.A.3c	Ratios and Proportional Relationships	Ratios and Proportional Relationships
3	Multiple Choice	1	CCSS.Math.Content.6.NS.B.4	The Number System	The Number System
4	Multiple Choice	1	CCSS.Math.Content.6.EE.C.9	Expressions and Equations	Expressions and Equations
5	Multiple Choice	1	CCSS.Math.Content.6.EE.A.2c	Expressions and Equations	Expressions and Equations
6	Multiple Choice	1	CCSS.Math.Content.6.RP.A.1	Ratios and Proportional Relationships	Ratios and Proportional Relationships
7	Multiple Choice	1	CCSS.Math.Content.6.G.A.3	Geometry	
8	Multiple Choice	1	CCSS.Math.Content.6.EE.B.7	Expressions and Equations	Expressions and Equations
9	Multiple Choice	1	CCSS.Math.Content.6.NS.C.6c	The Number System	The Number System

Question	Type	Points	Standard	Strand	Subscore
10	Multiple Choice	1	CCSS.Math.Content.6.G.A.2	Geometry	
11	Multiple Choice	1	CCSS.Math.Content.6.RP.A.3c	Ratios and Proportional Relationships	Ratios and Proportional Relationships
12	Multiple Choice	1	CCSS.Math.Content.6.EE.C.9	Expressions and Equations	Expressions and Equations
13	Multiple Choice	1	CCSS.Math.Content.6.EE.A.1	Expressions and Equations	Expressions and Equations
14	Multiple Choice	1	CCSS.Math.Content.6.RP.A.3b	Ratios and Proportional Relationships	Ratios and Proportional Relationships
15	Multiple Choice	1	CCSS.Math.Content.6.EE.B.5	Expressions and Equations	Expressions and Equations
16	Multiple Choice	1	CCSS.Math.Content.6.G.A.1	Geometry	
17	Multiple Choice	1	CCSS.Math.Content.6.NS.C.6	The Number System	The Number System
18	Multiple Choice	1	CCSS.Math.Content.6.RP.A.3b	Ratios and Proportional Relationships	Ratios and Proportional Relationships
19	Multiple Choice	1	CCSS.Math.Content.6.EE.A.3	Expressions and Equations	Expressions and Equations
20	Multiple Choice	1	CCSS.Math.Content.5.G.A.2	The Number System	The Number System
21	Multiple Choice	1	CCSS.Math.Content.6.G.A.4	Geometry	
22	Multiple Choice	1	CCSS.Math.Content.6.EE.A.4	Expressions and Equations	Expressions and Equations
23	Multiple Choice	1	CCSS.Math.Content.6.RP.A.3a	Ratios and Proportional Relationships	Ratios and Proportional Relationships
24	Multiple Choice	1	CCSS.Math.Content.6.EE.B.6	Expressions and Equations	Expressions and Equations
Session 2					
25	Multiple Choice	1	CCSS.Math.Content.6.NS.C.5	The Number System	The Number System
26	Multiple Choice	1	CCSS.Math.Content.6.RP.A.1	Ratios and Proportional Relationships	Ratios and Proportional Relationships
27	Multiple Choice	1	CCSS.Math.Content.6.EE.B.8	Expressions and Equations	Expressions and Equations
28	Multiple Choice	1	CCSS.Math.Content.6.G.A.3	Geometry	
29	Multiple Choice	1	CCSS.Math.Content.6.NS.A.1	The Number System	The Number System
30	Multiple Choice	1	CCSS.Math.Content.6.EE.A.3	Expressions and Equations	Expressions and Equations
31	Multiple Choice	1	CCSS.Math.Content.6.RP.A.2	Ratios and Proportional Relationships	Ratios and Proportional Relationships
32	Constructed Response	2	CCSS.Math.Content.6.NS.A.1	The Number System	The Number System

Question	Type	Points	Standard	Strand	Subscore
33	Constructed Response	2	CCSS.Math.Content.6.EE.A.2a	Expressions and Equations	Expressions and Equations
34	Constructed Response	2	CCSS.Math.Content.6.RP.A.3d	Ratios and Proportional Relationships	Ratios and Proportional Relationships
35	Constructed Response	2	CCSS.Math.Content.6.EE.A.1	Expressions and Equations	Expressions and Equations
36	Constructed Response	2	CCSS.Math.Content.6.NS.C.6b	The Number System	The Number System
37	Constructed Response	2	CCSS.Math.Content.6.RP.A.2	Ratios and Proportional Relationships	Ratios and Proportional Relationships
38	Constructed Response	2	CCSS.Math.Content.6.G.A.2	Geometry	
39	Constructed Response	3	CCSS.Math.Content.6.EE.B.7	Expressions and Equations	Expressions and Equations

*This item map is intended to identify the primary analytic skills necessary to successfully answer each question. However, some questions measure proficiencies described in multiple standards, including a balanced combination of procedural and conceptual understanding.

Table B.11. Mathematics Grade 7 Operational Item Map

Question	Type	Points	Standard	Cluster	Subscore
Session 1					
1	Multiple Choice	1	CCSS.Math.Content.7.EE.B.4a	Expressions and Equations	Expressions and Equations
2	Multiple Choice	1	CCSS.Math.Content.7.NS.A.2d	The Number System	The Number System
3	Multiple Choice	1	CCSS.Math.Content.7.RP.A.3	Ratios and Proportional Relationships	Ratios and Proportional Relationships
4	Multiple Choice	1	CCSS.Math.Content.7.G.B.4	Geometry	
5	Multiple Choice	1	CCSS.Math.Content.7.RP.A.1	Ratios and Proportional Relationships	Ratios and Proportional Relationships
6	Multiple Choice	1	CCSS.Math.Content.7.NS.A.3	The Number System	The Number System
7	Multiple Choice	1	CCSS.Math.Content.7.RP.A.2b	Ratios and Proportional Relationships	Ratios and Proportional Relationships
8	Multiple Choice	1	CCSS.Math.Content.7.NS.A.1a	The Number System	The Number System
9	Multiple Choice	1	CCSS.Math.Content.7.EE.B.3	Expressions and Equations	Expressions and Equations
10	Multiple Choice	1	CCSS.Math.Content.7.RP.A.3	Ratios and Proportional Relationships	Ratios and Proportional Relationships
11	Multiple Choice	1	CCSS.Math.Content.7.EE.A.2	Expressions and Equations	Expressions and Equations
12	Multiple Choice	1	CCSS.Math.Content.7.NS.A.3	The Number System	The Number System

Question	Type	Points	Standard	Cluster	Subscore
13	Multiple Choice	1	CCSS.Math.Content.7.RP.A.1	Ratios and Proportional Relationships	Ratios and Proportional Relationships
14	Multiple Choice	1	CCSS.Math.Content.7.RP.A.3	Ratios and Proportional Relationships	Ratios and Proportional Relationships
15	Multiple Choice	1	CCSS.Math.Content.7.RP.A.2a	Ratios and Proportional Relationships	Ratios and Proportional Relationships
16	Multiple Choice	1	CCSS.Math.Content.7.EE.A.1	Expressions and Equations	Expressions and Equations
17	Multiple Choice	1	CCSS.Math.Content.7.SP.C.7b	Statistics and Probability	
18	Multiple Choice	1	CCSS.Math.Content.7.EE.B.3	Expressions and Equations	Expressions and Equations
19	Multiple Choice	1	CCSS.Math.Content.7.SP.A.2	Statistics and Probability	
20	Multiple Choice	1	CCSS.Math.Content.7.G.A.1	Geometry	
21	Multiple Choice	1	CCSS.Math.Content.7.NS.A.1c	The Number System	The Number System
22	Multiple Choice	1	CCSS.Math.Content.7.EE.B.4b	Expressions and Equations	Expressions and Equations
23	Multiple Choice	1	CCSS.Math.Content.7.SP.A.1	Statistics and Probability	
24	Multiple Choice	1	CCSS.Math.Content.7.SP.C.6	Statistics and Probability	
25	Multiple Choice	1	CCSS.Math.Content.7.NS.A.1a	The Number System	The Number System
26	Multiple Choice	1	CCSS.Math.Content.7.EE.B.3	Expressions and Equations	Expressions and Equations
Session 2					
27	Multiple Choice	1	CCSS.Math.Content.7.RP.A.2c	Ratios and Proportional Relationships	Ratios and Proportional Relationships
28	Multiple Choice	1	CCSS.Math.Content.7.EE.A.1	Expressions and Equations	Expressions and Equations
29	Multiple Choice	1	CCSS.Math.Content.6.SP.B.5c	Statistics and Probability	
30	Multiple Choice	1	CCSS.Math.Content.7.SP.C.6	Statistics and Probability	
31	Multiple Choice	1	CCSS.Math.Content.7.EE.B.4b	Expressions and Equations	Expressions and Equations
32	Multiple Choice	1	CCSS.Math.Content.7.NS.A.2c	The Number System	The Number System
33	Multiple Choice	1	CCSS.Math.Content.7.EE.B.4a	Expressions and Equations	Expressions and Equations
34	Constructed Response	2	CCSS.Math.Content.7.EE.B.3	Expressions and Equations	Expressions and Equations
35	Constructed Response	2	CCSS.Math.Content.7.RP.A.3	Ratios and Proportional Relationships	Ratios and Proportional Relationships

Question	Type	Points	Standard	Cluster	Subscore
36	Constructed Response	2	CCSS.Math.Content.7.RP.A.1	Ratios and Proportional Relationships	Ratios and Proportional Relationships
37	Constructed Response	2	CCSS.Math.Content.7.EE.B.4a	Expressions and Equations	Expressions and Equations
38	Constructed Response	2	CCSS.Math.Content.7.SP.A.2	Statistics and Probability	
39	Constructed Response	2	CCSS.Math.Content.7.NS.A.3	The Number System	The Number System
40	Constructed Response	2	CCSS.Math.Content.7.RP.A.2c	Ratios and Proportional Relationships	Ratios and Proportional Relationships
41	Constructed Response	3	CCSS.Math.Content.7.EE.B.3	Expressions and Equations	Expressions and Equations

*This item map is intended to identify the primary analytic skills necessary to successfully answer each question. However, some questions measure proficiencies described in multiple standards, including a balanced combination of procedural and conceptual understanding.

Table B.12. Mathematics Grade 8 Operational Item Map

Question	Type	Points	Standard	Cluster	Subscore
Session 1					
1	Multiple Choice	1	CCSS.Math.Content.8.F.A.3	Functions	Functions
2	Multiple Choice	1	CCSS.Math.Content.8.EE.A.4	Expressions and Equations	Expressions and Equations
3	Multiple Choice	1	CCSS.Math.Content.8.F.B.4	Functions	Functions
4	Multiple Choice	1	CCSS.Math.Content.8.G.C.9	Geometry	Geometry
5	Multiple Choice	1	CCSS.Math.Content.8.EE.B.5	Expressions and Equations	Expressions and Equations
6	Multiple Choice	1	CCSS.Math.Content.8.F.A.3	Functions	Functions
7	Multiple Choice	1	CCSS.Math.Content.8.EE.B.6	Expressions and Equations	Expressions and Equations
8	Multiple Choice	1	CCSS.Math.Content.8.EE.A.4	Expressions and Equations	Expressions and Equations
9	Multiple Choice	1	CCSS.Math.Content.8.G.A.2	Geometry	Geometry
10	Multiple Choice	1	CCSS.Math.Content.8.EE.C.8b	Expressions and Equations	Expressions and Equations
11	Multiple Choice	1	CCSS.Math.Content.8.EE.C.8b	Expressions and Equations	Expressions and Equations
12	Multiple Choice	1	CCSS.Math.Content.8.F.B.4	Functions	Functions
13	Multiple Choice	1	CCSS.Math.Content.8.G.A.4	Geometry	Geometry

Question	Type	Points	Standard	Cluster	Subscore
14	Multiple Choice	1	CCSS.Math.Content.7.G.B.6	Geometry	Geometry
15	Multiple Choice	1	CCSS.Math.Content.8.EE.B.6	Expressions and Equations	Expressions and Equations
16	Multiple Choice	1	CCSS.Math.Content.8.EE.C.7a	Expressions and Equations	Expressions and Equations
17	Multiple Choice	1	CCSS.Math.Content.8.SP.A.3	Statistics and Probability	
18	Multiple Choice	1	CCSS.Math.Content.8.SP.A.2	Statistics and Probability	
19	Multiple Choice	1	CCSS.Math.Content.8.EE.B.5	Expressions and Equations	Expressions and Equations
20	Multiple Choice	1	CCSS.Math.Content.8.G.C.9	Geometry	Geometry
21	Multiple Choice	1	CCSS.Math.Content.8.EE.A.3	Expressions and Equations	Expressions and Equations
22	Multiple Choice	1	CCSS.Math.Content.8.EE.A.1	Expressions and Equations	Expressions and Equations
23	Multiple Choice	1	CCSS.Math.Content.8.F.A.2	Functions	Functions
24	Multiple Choice	1	CCSS.Math.Content.8.SP.A.1	Statistics and Probability	
25	Multiple Choice	1	CCSS.Math.Content.8.G.A.5	Geometry	Geometry
26	Multiple Choice	1	CCSS.Math.Content.8.F.B.5	Functions	Functions
Session 2					
27	Multiple Choice	1	CCSS.Math.Content.8.F.A.1	Functions	Functions
28	Multiple Choice	1	CCSS.Math.Content.8.G.C.9	Geometry	Geometry
29	Multiple Choice	1	CCSS.Math.Content.8.SP.A.2	Statistics and Probability	
30	Multiple Choice	1	CCSS.Math.Content.8.EE.A.1	Expressions and Equations	Expressions and Equations
31	Multiple Choice	1	CCSS.Math.Content.8.SP.A.1	Statistics and Probability	
32	Multiple Choice	1	CCSS.Math.Content.8.G.A.3	Geometry	Geometry
33	Multiple Choice	1	CCSS.Math.Content.8.F.A.1	Functions	Functions
34	Constructed Response	2	CCSS.Math.Content.8.EE.A.4	Expressions and Equations	Expressions and Equations
35	Constructed Response	2	CCSS.Math.Content.8.F.A.3	Functions	Functions
36	Constructed Response	2	CCSS.Math.Content.8.EE.C.8b	Expressions and Equations	Expressions and Equations

Question	Type	Points	Standard	Cluster	Subscore
37	Constructed Response	2	CCSS.Math.Content.8.F.B.4	Functions	Functions
38	Constructed Response	2	CCSS.Math.Content.8.EE.B.5	Expressions and Equations	Expressions and Equations
39	Constructed Response	2	CCSS.Math.Content.8.G.A.3	Geometry	Geometry
40	Constructed Response	2	CCSS.Math.Content.8.G.C.9	Geometry	Geometry
41	Constructed Response	3	CCSS.Math.Content.8.F.A.3	Functions	Functions

*This item map is intended to identify the primary analytic skills necessary to successfully answer each question. However, some questions measure proficiencies described in multiple standards, including a balanced combination of procedural and conceptual understanding.

Appendix C: Raw-to-Scale Score Tables

Tables C.1–C.12 show the PBT raw-to-scale score conversion tables, while Tables C.13–C.24 show the CBT raw-to-scale score conversion tables.

Table C.1. PBT ELA Grade 3 RSSS Table

Raw Score	Scale Score	CSEM
0	542	28
1	547	23
2	552	20
3	556	17
4	561	15
5	565	13
6	573	11
7	583	10
8	585	10
9	590	9
10	595	9
11	602	8
12	604	8
13	608	8
14	613	8
15	619	9
16	629	11
17	637	15
18	642	17

Table C.2. PBT ELA Grade 4 RSSS Table

Raw Score	Scale Score	CSEM
0	542	34
1	547	29
2	552	25
3	556	21
4	561	18
5	565	16
6	573	13
7	584	11
8	585	10

Raw Score	Scale Score	CSEM
9	590	10
10	595	9
11	603	9
12	604	10
13	610	10
14	619	11
15	624	12
16	634	15
17	650	22
18	655	24

Table C.3. PBT ELA Grade 5 RSSS Table

Raw Score	Scale Score	CSEM
0	523	65
1	528	56
2	532	48
3	537	41
4	542	35
5	546	29
6	551	24
7	555	20
8	564	14
9	570	11
10	574	10
11	578	8
12	581	8
13	584	7
14	587	7
15	590	6
16	594	6
17	595	6
18	598	6
19	601	6
20	604	7
21	609	7
22	611	7

Raw Score	Scale Score	CSEM
23	615	8
24	622	9
25	627	10
26	636	13
27	651	21
28	656	23

Table C.4. PBT ELA Grade 6 RSSS Table

Raw Score	Scale Score	CSEM
0	523	37
1	527	33
2	532	29
3	537	25
4	541	22
5	546	19
6	550	17
7	555	15
8	562	13
9	568	11
10	573	10
11	577	9
12	581	8
13	585	8
14	590	7
15	591	7
16	594	7
17	597	7
18	602	7
19	603	7
20	606	7
21	610	8
22	614	8
23	619	9
24	625	10
25	632	12
26	643	15

Raw Score	Scale Score	CSEM
27	647	17
28	652	19

Table C.5. PBT ELA Grade 7 RSSS Table

Raw Score	Scale Score	CSEM
0	523	39
1	527	36
2	532	33
3	536	30
4	541	28
5	545	26
6	550	23
7	555	21
8	559	19
9	568	15
10	575	13
11	580	11
12	585	10
13	591	9
14	593	8
15	596	8
16	599	7
17	602	7
18	607	7
19	608	7
20	612	7
21	615	7
22	618	7
23	623	7
24	627	8
25	633	10
26	641	13
27	646	15
28	650	17

Table C.6. PBT ELA Grade 8 RSSS Table

Raw Score	Scale Score	CSEM
0	523	91
1	528	74
2	533	59
3	537	47
4	542	37
5	546	30
6	551	23
7	555	19
8	560	15
9	567	12
10	572	10
11	576	9
12	579	8
13	584	8
14	586	7
15	589	7
16	592	7
17	595	7
18	598	7
19	603	7
20	604	7
21	607	7
22	611	7
23	617	7
24	619	8
25	624	9
26	631	10
27	642	15
28	647	18

Table C.7. PBT Mathematics Grade 3 RSSS Table

Raw Score	Scale Score	CSEM
0	541	41
1	546	31
2	550	23

Raw Score	Scale Score	CSEM
3	555	18
4	560	15
5	568	11
6	574	9
7	578	8
8	583	7
9	587	7
10	590	7
11	593	7
12	600	7
13	601	7
14	605	7
15	610	8
16	615	9
17	622	10
18	633	12
19	637	15

Table C.8. PBT Mathematics Grade 4 RSSS Table

Raw Score	Scale Score	CSEM
0	536	105
1	541	86
2	546	67
3	550	51
4	555	37
5	559	27
6	571	13
7	577	9
8	581	8
9	585	7
10	588	6
11	591	6
12	594	5
13	596	5
14	599	5
15	602	5

Raw Score	Scale Score	CSEM
16	604	5
17	606	5
18	609	5
19	614	6
20	616	6
21	621	7
22	630	10
23	634	13

Table C.9. PBT Mathematics Grade 5 RSSS Table

Raw Score	Scale Score	CSEM
0	541	54
1	545	42
2	550	32
3	554	24
4	559	19
5	564	14
6	572	10
7	577	8
8	581	8
9	585	7
10	589	6
11	592	6
12	595	6
13	597	6
14	600	6
15	604	5
16	605	5
17	608	5
18	611	5
19	616	5
20	618	6
21	622	7
22	630	10
23	634	13

Table C.10. PBT Mathematics Grade 6 RSSS Table

Raw Score	Scale Score	CSEM
0	546	43
1	551	35
2	555	29
3	560	24
4	564	20
5	569	16
6	577	11
7	582	8
8	586	7
9	589	6
10	592	6
11	594	5
12	597	5
13	599	5
14	601	5
15	604	5
16	606	5
17	608	5
18	610	5
19	613	5
20	616	5
21	619	6
22	624	7
23	631	10
24	636	12

Table C.11. PBT Mathematics Grade 7 RSSS Table

Raw Score	Scale Score	CSEM
0	529	428
1	534	296
2	538	205
3	543	142
4	548	99
5	552	70
6	557	49

Raw Score	Scale Score	CSEM
7	575	15
8	583	11
9	588	9
10	593	7
11	595	6
12	597	5
13	600	5
14	602	4
15	604	4
16	606	4
17	607	4
18	609	4
19	611	4
20	613	4
21	615	4
22	618	5
23	621	5
24	625	6
25	633	10
26	637	13

Table C.12. PBT Mathematics Grade 8 RSSS Table

Raw Score	Scale Score	CSEM
0	544	79
1	548	67
2	553	58
3	557	50
4	562	43
5	567	37
6	571	30
7	576	24
8	585	13
9	590	9
10	596	7
11	597	6
12	599	5

Raw Score	Scale Score	CSEM
13	602	5
14	604	5
15	606	5
16	608	4
17	610	4
18	612	4
19	614	4
20	616	4
21	618	5
22	622	5
23	624	6
24	629	7
25	636	10
26	641	13

Table C.13. CBT ELA Grade 3 RSSS Table

Raw Score	Scale Score*	CSEM
0	544	26
1	549	22
2	554	18
3	558	16
4	563	14
5	567	13
6	575	11
7	585	10
8	587	9
9	592	9
10	597	8
11	604	8
12	606	8
13	610	8
14	615	8
15	621	9
16	631	12
17	639	15
18	642	17

* A CBT mode adjustment has been taken into account for these scale scores

Table C.14. CBT ELA Grade 4 RSSS Table

Raw Score	Scale Score*	CSEM
0	544	32
1	549	27
2	554	23
3	558	20
4	563	17
5	567	15
6	575	12
7	586	10
8	587	10
9	592	9
10	597	9
11	605	10

Raw Score	Scale Score*	CSEM
12	606	10
13	612	10
14	621	12
15	626	13
16	636	16
17	652	23
18	655	24

* A CBT mode adjustment has been taken into account for these scale scores

Table C.15. CBT ELA Grade 5 RSSS Table

Raw Score	Scale Score*	CSEM
0	525	62
1	530	52
2	534	46
3	539	38
4	544	32
5	548	27
6	553	22
7	557	18
8	566	13
9	572	10
10	576	9
11	580	8
12	583	7
13	586	7
14	589	6
15	592	6
16	596	6
17	597	6
18	600	6
19	603	7
20	606	7
21	611	7
22	613	8
23	617	8
24	624	10

Raw Score	Scale Score*	CSEM
25	629	11
26	638	14
27	653	22
28	656	23

* A CBT mode adjustment has been taken into account for these scale scores

Table C.16. CBT ELA Grade 6 RSSS Table

Raw Score	Scale Score*	CSEM
0	524	36
1	528	32
2	533	28
3	538	24
4	542	22
5	547	19
6	551	17
7	556	15
8	563	13
9	569	11
10	574	10
11	578	9
12	582	8
13	586	8
14	591	7
15	592	7
16	595	7
17	598	7
18	603	7
19	604	7
20	607	7
21	611	8
22	615	8
23	620	9
24	626	10
25	633	12
26	644	16
27	648	18

Raw Score	Scale Score*	CSEM
28	652	19

* A CBT mode adjustment has been taken into account for these scale scores

Table C.17. CBT ELA Grade 7 RSSS Table

Raw Score	Scale Score*	CSEM
0	525	37
1	529	34
2	534	31
3	538	29
4	543	27
5	547	25
6	552	23
7	557	20
8	561	19
9	570	15
10	577	12
11	582	11
12	587	9
13	593	8
14	595	8
15	598	7
16	601	7
17	604	7
18	609	7
19	610	7
20	614	7
21	617	7
22	620	7
23	625	8
24	629	9
25	635	10
26	643	13
27	648	16
28	650	17

* A CBT mode adjustment has been taken into account for these scale scores

Table C.18. CBT ELA Grade 8 RSSS Table

Raw Score	Scale Score*	CSEM
0	525	85
1	530	67
2	535	52
3	539	43
4	544	33
5	548	27
6	553	21
7	557	18
8	562	14
9	569	11
10	574	9
11	578	8
12	581	8
13	586	7
14	588	7
15	591	7
16	594	7
17	597	7
18	600	7
19	605	7
20	606	7
21	609	7
22	613	7
23	619	8
24	621	8
25	626	9
26	633	11
27	644	16
28	647	18

* A CBT mode adjustment has been taken into account for these scale scores

Table C.19. CBT Mathematics Grade 3 RSSS Table

Raw Score	Scale Score*	CSEM
0	542	39
1	547	29
2	551	23
3	556	17
4	561	14
5	569	10
6	575	9
7	579	8
8	584	7
9	588	7
10	591	7
11	594	7
12	601	7
13	602	7
14	606	7
15	611	8
16	616	9
17	623	10
18	634	13
19	637	15

* A CBT mode adjustment has been taken into account for these scale scores

Table C.20. CBT Mathematics Grade 4 RSSS Table

Raw Score	Scale Score*	CSEM
0	537	102
1	542	82
2	547	62
3	551	48
4	556	34
5	560	26
6	572	12
7	578	9
8	582	7
9	586	7
10	589	6

Raw Score	Scale Score*	CSEM
11	592	6
12	595	5
13	597	5
14	600	5
15	603	5
16	605	5
17	607	5
18	610	5
19	615	6
20	617	6
21	622	7
22	631	11
23	634	13

* A CBT mode adjustment has been taken into account for these scale scores

Table C.21. CBT Mathematics Grade 5 RSSS Table

Raw Score	Scale Score*	CSEM
0	542	50
1	546	40
2	551	30
3	555	23
4	560	17
5	565	13
6	573	10
7	578	8
8	582	7
9	586	7
10	590	6
11	593	6
12	596	6
13	598	6
14	601	5
15	605	5
16	606	5
17	609	5
18	612	5

Raw Score	Scale Score*	CSEM
19	617	6
20	619	6
21	623	7
22	631	10
23	634	13

* A CBT mode adjustment has been taken into account for these scale scores

Table C.22. CBT Mathematics Grade 6 RSSS Table

Raw Score	Scale Score*	CSEM
0	547	41
1	552	33
2	556	28
3	561	23
4	565	19
5	570	15
6	578	10
7	583	8
8	587	7
9	590	6
10	593	5
11	595	5
12	598	5
13	600	5
14	602	5
15	605	5
16	607	5
17	609	5
18	611	5
19	614	5
20	617	5
21	620	6
22	625	7
23	632	10
24	636	12

* A CBT mode adjustment has been taken into account for these scale scores

Table C.23. CBT Mathematics Grade 7 RSSS Table

Raw Score	Scale Score*	CSEM
0	531	370
1	536	249
2	540	181
3	545	122
4	550	83
5	554	62
6	559	43
7	577	14
8	585	10
9	590	8
10	595	6
11	597	5
12	599	5
13	602	4
14	604	4
15	606	4
16	608	4
17	609	4
18	611	4
19	613	4
20	615	4
21	617	4
22	620	5
23	623	6
24	627	7
25	635	11
26	637	13

* A CBT mode adjustment has been taken into account for these scale scores

Table C.24. CBT Mathematics Grade 8 RSSS Table

Raw Score	Scale Score*	CSEM
0	545	76
1	549	66
2	554	56
3	558	49
4	563	42
5	568	35
6	572	29
7	577	22
8	586	12
9	591	9
10	597	6
11	598	6
12	600	5
13	603	5
14	605	5
15	607	4
16	609	4
17	611	4
18	613	4
19	615	4
20	617	5
21	619	5
22	623	5
23	625	6
24	630	7
25	637	10
26	641	13

*A CBT mode adjustment has been taken into account for these scale scores

Appendix D: Study of Operational Test Mode Comparability

Section D.1. Introduction

Being mindful of the variability in the manner in which instruction was being provided to students across the State in response to COVID-19, NYSED decided to administer the pre-existing test forms from one of the last two administrations (either 2018 or 2019) for the 2021 administration of Grades 3–8 ELA and Mathematics Tests. Schools were only required to administer Session 1 of the ELA and Mathematics Tests to students, and Session 1 tests were reduced in length by the removal of the embedded multiple-choice test questions.

A new raw-score-to-scale-score (RSSS) table for each grade/subject was generated for the shortened test based on Session 1 items using the inverse-TCC method with the most recent banked item IRT parameters. Tables D.1 and D.2 summarize the number of items by item type for each session for the ELA and Math test forms.

Table D.1. Summary of Item Type by Session on 2021 Grades 3–8 ELA Test Forms

Subject	Grade	Session	Strand	# Items			Raw Score Points
				MC	CR2	CR4	
ELA	3	1	Reading	18	0	0	18
		2	Writing	0	6	1	16
	4	1	Reading	18	0	0	18
		2	Writing	0	6	1	16
	5	1	Reading	28	0	0	28
		2	Writing	0	6	1	16
	6	1	Reading	28	0	0	28
		2	Writing	0	6	1	16
	7	1	Reading	28	0	0	28
		2	Writing	0	7	1	18
	8	1	Reading	28	0	0	28
		2	Writing	0	7	1	18

Table D.2. Summary of Item Type on 2021 Grades 3–8 Math Test Forms

Subject	Grade	Session	Calculator	# Items			Raw Score Points
				MC	CR2	CR3	
Math	3	1	No	19	0	0	19
		2	No	8	6	1	23
	4	1	No	23	0	0	23
		2	No	8	6	1	23
	5	1	No	23	0	0	23
		2	No	8	6	1	23
	6	1	No	24	0	0	24
		2	Yes	7	7	1	24

Subject	Grade	Session	Calculator	# Items			Raw Score Points
				MC	CR2	CR3	
Math	7	1	Yes	26	0	0	26
		2	Yes	7	7	1	24
	8	1	Yes	26	0	0	26
		2	Yes	7	7	1	24

In the 2021 administration of Grades 3–8 ELA and Mathematics Tests, NYSED continued to offer its operational test in a computer-based testing (CBT) environment for the Grades 3–8 ELA and Mathematics tests. The schools have the option to administer the tests via paper-based testing (PBT) or computer-based testing (CBT). Because of the test structure change, the adjustment on the CBT mode RSSS tables needs to be reevaluated. In response to COVID-19, NYSED provided schools much more flexibility than had been necessary in past years in the local in-school scheduling and administration of the Spring 2021 ELA and Mathematics Tests. Given these unique circumstances, schools were not expected to bring students into the building to participate in the tests if the students were receiving entirely remote instruction during the testing windows. As a result, the population of students testing in 2021 was not representative of the student population overall. Therefore, the matched samples in the mode comparability studies conducted in the 2018 and 2019 administrations were used to check whether there is a difference in performance between the test modes based on Session 1 performance.

Section D.2. Method

This section outlines the steps needed for reevaluating the CBT mode adjustment in this mode comparability study.

1) Determine the samples used to reevaluate the adjustment of the student scores on the CBT test mode. A propensity score matching approach has been implemented in the 2018 and 2019 administrations of Grades 3–8 ELA and Mathematics Tests to generate the CBT and PBT samples that were comparable on selected covariates that may affect student performance, aside from the test mode itself. The covariate balance between the CBT and PBT samples was evaluated after propensity score matching, which showed well-matched PBT and CBT groups were achieved. (Please see the "Comparability of Spring 2018 Grades 3-8 English Language Arts and Mathematics Paper-based and Computer-based Tests" or the "Comparability of Spring 2019 Grades 3-8 English Language Arts and Mathematics Paper-based and Computer-based Tests" for more information.) These matched samples were used to evaluate test-level mode comparability in the 2018 and 2019 administrations. Therefore, the same matched samples can be used to evaluate test-level mode comparability based on Session 1 item responses. For each subject and grade level, the samples used for this study were from the corresponding year when the test forms were originally administered.
2) The distributions of raw scores on Session 1 for the matched PBT and CBT samples are first reviewed. This enables a direct means of detecting possible mode effects on Session 1 of the test.
3) Next, the distributions of scale scores for the matched PBT and CBT samples are reviewed. The scale scores are derived using the new operational raw-score-to-scale-score (RSSS) tables which are generated based on Session 1 items. The mode treatment effect is calculated as the difference in scale score means for the matched PBT and CBT samples.
4) Calculate the difference on the selected scores for the matched samples. CBT can be considered the "treatment" and PBT can be considered the "control" condition. Both difference (**Δ**) and standardized difference (*d*) between CBT (*t*) and PBT (*c*) for both before (*b*) and after matching (*m*) are calculated as below:

$$\mathbf{\Delta}_{bk} = \bar{x}_{tk} - \bar{x}_{ck} \quad (1)$$

$$d_{bk} = (\bar{x}_{tk} - \bar{x}_{ck}) \Big/ \sqrt{(s_{tk}^2 + s_{ck}^2)/2} \quad (2)$$

$$\mathbf{\Delta}_{mk} = \bar{x}_{tmk} - \bar{x}_{cmk} \quad (3)$$

$$d_{mk} = (\bar{x}_{tmk} - \bar{x}_{cmk}) \Big/ \sqrt{(s_{tk}^2 + s_{ck}^2)/2} \quad (4)$$

Section D.3. Results

Questar calculated the sample means for each matched sample and their standardized differences before and after matching for the following variables:

- 2017 or 2018 Scale Score (SS): the prior year (grade n – 1) scale score, which was the proxy for prior ability that was entered as a key predictor into the propensity score matching model;
- 2018 or 2019 Raw Score (RS): the current year operational raw score for the original test;
- 2018 or 2019 Scale Score (SS): the current year scale score for the original test;
- 2018 or 2019 Raw Score on Session 1 (Session 1 RS): the current year operational raw score on Session 1;
- 2018 or 2019 Scale Score based on Session 1 (Session 1 SS): the current year scale score based on Session 1 RSSS table.

The test-level performance before and after matching is summarized in Tables D.3–D.6. The results for CBT scale score adjustment after matching comparison between the original test (Session 1 + 2) and Session 1 only are summarized in Table D.7.

Table D.3. Test-level Performance between Test Modes Before Matching – ELA

Test	Variable	PBT			CBT			Delta	*d*
		N	Mean	SD	N	Mean	SD		
ELA4	2017 SS	146350	309.65	34.44	14581	307.80	31.94	-1.85	-0.06
	2018 RS	165390	19.57	7.09	16282	18.34	6.60	-1.23	-0.18
	2018 SS	165390	600.26	20.16	16282	596.77	18.21	-3.50	-0.18
	2018 RS_Session 1	165390	10.98	3.92	16282	10.66	3.74	-0.32	-0.08
	2018 SS_Session 1	165390	602.36	24.01	16282	600.24	22.52	-2.12	-0.09
ELA5	2018 SS	125862	601.05	20.31	30142	599.13	18.34	-1.92	-0.10
	2019 RS	141544	27.33	9.02	33839	26.22	8.63	-1.11	-0.13
	2019 SS	141544	599.68	21.45	33839	597.02	19.89	-2.66	-0.13
	2019 RS_Session 1	141544	17.88	5.98	33839	17.56	5.87	-0.32	-0.05
	2019 SS_Session 1	141544	599.88	23.29	33839	598.68	22.59	-1.20	-0.05
ELA6	2017 SS	132281	303.23	38.68	16385	302.36	35.46	-0.87	-0.02
	2018 RS	151027	28.00	8.75	18988	27.22	8.27	-0.78	-0.09
	2018 SS	151027	600.25	20.19	18988	598.27	18.39	-1.98	-0.10
	2018 RS_Session 1	151027	17.72	5.57	18988	17.65	5.26	-0.07	-0.01
	2018 SS_Session 1	151027	600.50	22.12	18988	600.15	20.50	-0.35	-0.02
ELA7	2017 SS	125455	300.45	35.89	13900	301.77	33.82	1.32	0.04
	2018 RS	140012	28.98	9.20	15907	27.66	8.60	-1.32	-0.15
	2018 SS	140012	600.21	20.18	15907	597.27	18.18	-2.94	-0.15
	2018 RS_Session 1	140012	16.68	5.55	15907	16.25	5.22	-0.43	-0.08
	2018 SS_Session 1	140012	600.10	22.44	15907	598.50	20.94	-1.60	-0.07
ELA8	2018 SS	113222	601.24	19.82	22258	599.29	18.64	-1.95	-0.10
	2019 RS	126088	31.67	8.98	25878	30.27	8.89	-1.40	-0.16

Test	Variable	PBT			CBT			Delta	*d*
		N	Mean	SD	N	Mean	SD		
ELA8	2019 SS	126088	600.48	19.85	25878	597.22	18.92	-3.26	-0.17
	2019 RS_Session 1	126088	18.46	5.77	25878	18.04	5.66	-0.42	-0.07
	2019 SS_Session 1	126088	600.73	21.83	25878	599.10	21.18	-1.63	-0.08

Table D.4. Test-level Performance between Test Modes After Matching – ELA

Test	Variable	PBT			CBT			Delta	*d*
		N	Mean	SD	N	Mean	SD		
ELA4	2017 SS	11539	310.21	32.36	11539	311.41	30.69	1.20	0.04
	2018 RS	11539	19.39	6.67	11539	19.14	6.32	-0.25	-0.04
	2018 SS	11539	599.63	18.64	11539	598.99	17.34	-0.64	-0.03
	2018 RS_Session 1	11539	11.32	3.78	11539	11.06	3.63	-0.26	**-0.07**
	2018 SS_Session 1	11539	604.21	23.30	11539	602.55	22.07	-1.66	**-0.07**
ELA5	2018 SS	25210	599.85	19.49	25210	599.17	18.43	-0.68	-0.04
	2019 RS	25210	27.50	8.67	25210	26.53	8.60	-0.97	**-0.11**
	2019 SS	25210	600.01	20.41	25210	597.76	19.85	-2.25	**-0.11**
	2019 RS_Session 1	25210	18.35	5.76	25210	17.71	5.85	-0.64	**-0.11**
	2019 SS_Session 1	25210	601.60	22.45	25210	599.26	22.56	-2.34	**-0.10**
ELA6	2017 SS	14647	305.66	36.41	14647	305.63	34.04	-0.03	0.00
	2018 RS	14647	28.58	8.17	14647	28.27	7.84	-0.30	-0.04
	2018 SS	14647	601.35	18.57	14647	600.57	17.50	-0.78	-0.04
	2018 RS_Session 1	14647	18.57	5.15	14647	18.25	5.02	-0.32	**-0.06**
	2018 SS_Session 1	14647	603.75	20.44	14647	602.44	19.60	-1.31	**-0.06**
ELA7	2017 SS	11736	304.22	34.24	11736	305.05	32.26	0.83	0.02
	2018 RS	11736	29.30	8.78	11736	28.86	8.18	-0.44	-0.05
	2018 SS	11736	600.77	19.03	11736	599.82	17.19	-0.95	-0.05
	2018 RS_Session 1	11736	17.30	5.31	11736	16.86	5.08	-0.44	**-0.08**
	2018 SS_Session 1	11736	602.54	21.22	11736	601.00	20.09	-1.54	**-0.07**
ELA8	2018 SS	19875	599.02	19.83	19875	599.08	18.72	0.06	0.00
	2019 RS	19875	31.00	9.18	19875	30.63	8.74	-0.37	-0.04
	2019 SS	19875	598.94	20.01	19875	597.99	18.61	-0.95	-0.05
	2019 RS_Session 1	19875	18.69	5.70	19875	18.22	5.61	-0.47	**-0.08**
	2019 SS_Session 1	19875	601.51	21.49	19875	599.76	20.97	-1.75	**-0.08**

Note. The standardized difference (*d*) with an absolute value greater than 0.05 after matching was bolded.

Table D.5. Test-level Performance between Test Modes Before Matching – Math

Test	Variable	PBT			CBT			Delta	*d*
		N	Mean	SD	N	Mean	SD		
Math4	2018 SS	140255	600.88	19.98	18793	598.97	18.93	-1.91	-0.10
	2019 RS	161135	28.46	11.12	21198	27.28	10.52	-1.18	-0.11
	2019 SS	161135	600.58	20.42	21198	598.10	18.57	-2.48	-0.13
	2019 RS_Session 1	161135	14.54	5.63	21198	14.21	5.41	-0.33	-0.06

Test	Variable	PBT			CBT			Delta	*d*
		N	Mean	SD	N	Mean	SD		
Math4	2019 SS_Session 1	161135	600.30	20.09	21198	599.11	19.10	-1.19	-0.06
Math5	2018 SS	132758	601.51	19.83	19668	600.07	18.23	-1.44	-0.08
	2019 RS	148983	27.67	11.54	21975	25.85	11.03	-1.82	-0.16
	2019 SS	148983	601.08	20.48	21975	597.60	19.08	-3.48	-0.18
	2019 RS_Session 1	148983	14.30	5.65	21975	13.77	5.46	-0.53	-0.10
	2019 SS_Session 1	148983	600.78	19.92	21975	598.86	19.18	-1.92	-0.10
Math6	2017 SS	132143	309.69	38.51	11735	312.37	34.21	2.68	0.07
	2018 RS	150945	23.55	11.60	13484	23.67	10.41	0.12	0.01
	2018 SS	150945	599.92	20.28	13484	600.46	17.27	0.54	0.03
	2018 RS_Session 1	150945	13.86	6.21	13484	14.06	5.64	0.20	0.03
	2018 SS_Session 1	150945	600.24	19.36	13484	600.93	17.01	0.69	0.04
Math7	2018 SS	121675	600.79	19.97	20101	602.10	18.17	1.31	0.07
	2019 RS	136764	28.58	13.50	23151	26.93	12.62	-1.65	-0.13
	2019 SS	136764	601.93	20.36	23151	599.35	18.73	-2.58	-0.13
	2019 RS_Session 1	136764	15.19	6.51	23151	14.58	6.06	-0.61	-0.10
	2019 SS_Session 1	136764	601.05	21.61	23151	599.48	19.97	-1.57	-0.08
Math8	2018 SS	82260	597.22	19.31	11839	593.92	17.60	-3.30	-0.18
	2019 RS	93095	25.89	12.67	13994	22.56	11.16	-3.33	-0.28
	2019 SS	93095	601.02	21.32	13994	595.62	19.68	-5.40	-0.26
	2019 RS_Session 1	93095	14.18	6.33	13994	12.78	5.57	-1.40	-0.23
	2019 SS_Session 1	93095	601.69	19.61	13994	597.63	17.57	-4.06	-0.22

Table D.6. Test-level Performance between Test Modes After Matching – Math

Test	Variable	PBT			CBT			Delta	*d*
		N	Mean	SD	N	Mean	SD		
Math4	2018 SS	16520	599.27	19.38	16520	598.88	18.97	-0.39	-0.02
	2019 RS	16520	28.18	10.72	16520	27.48	10.49	-0.70	**-0.06**
	2019 SS	16520	599.74	19.31	16520	598.43	18.49	-1.31	**-0.07**
	2019 RS_Session 1	16520	14.50	5.50	16520	14.29	5.41	-0.21	-0.04
	2019 SS_Session 1	16520	600.02	19.51	16520	599.39	19.10	-0.63	-0.03
Math5	2018 SS	17660	600.45	18.06	17660	600.63	18.02	0.18	0.01
	2019 RS	17660	27.62	10.91	17660	26.62	10.84	-1.00	**-0.09**
	2019 SS	17660	600.67	18.74	17660	598.98	18.53	-1.69	**-0.09**
	2019 RS_Session 1	17660	14.37	5.42	17660	14.13	5.38	-0.24	-0.04
	2019 SS_Session 1	17660	600.95	18.94	17660	600.16	18.75	-0.79	-0.04
Math6	2017 SS	9840	315.41	34.04	9840	314.89	33.00	-0.53	-0.01
	2018 RS	9840	25.97	10.68	9840	24.87	10.29	-1.10	**-0.10**
	2018 SS	9840	604.07	17.45	9840	602.47	16.65	-1.60	**-0.08**
	2018 RS_Session 1	9840	15.15	5.71	9840	14.67	5.54	-0.48	**-0.08**
	2018 SS_Session 1	9840	604.09	17.31	9840	602.74	16.56	-1.35	**-0.07**

Test	Variable	PBT			CBT			Delta	*d*
		N	Mean	SD	N	Mean	SD		
Math7	2018 SS	18446	602.15	18.56	18446	602.05	18.00	-0.10	-0.01
	2019 RS	18446	29.21	12.77	18446	27.61	12.53	-1.60	**-0.12**
	2019 SS	18446	602.61	18.59	18446	600.38	18.32	-2.24	**-0.11**
	2019 RS_Session 1	18446	15.53	6.18	18446	14.89	6.03	-0.64	**-0.10**
	2019 SS_Session 1	18446	602.22	19.91	18446	600.48	19.58	-1.74	**-0.08**
Math8	2018 SS	10133	595.08	17.56	10133	594.99	16.90	-0.09	0.00
	2019 RS	10133	24.81	11.49	10133	23.69	10.94	-1.12	**-0.09**
	2019 SS	10133	599.27	19.22	10133	597.74	18.54	-1.53	**-0.07**
	2019 RS_Session 1	10133	13.59	5.80	10133	13.28	5.51	-0.31	-0.05
	2019 SS_Session 1	10133	599.98	17.99	10133	599.26	17.03	-0.72	-0.04

Note. The standardized difference (*d*) with an absolute value greater than 0.05 after matching was bolded.

Section D.4. Discussion and Conclusions

D.4.1. Discussion

Based on the analyses described above, NYSED — in consultation with New York State's Assessment TAC and Questar — decided to apply an additive adjustment to CBT students' scale scores because it best balanced concerns about fairness, interpretability and face validity. NYSED also chose to set a ceiling above which the CBT students' scale scores would not be adjusted — namely the maximum observed scale score available to PBT students. In other words, the highest scale score on CBT was constrained to be equal to the highest scale score for PBT students.

The differences in the scale score means between the matched samples using the new RSSS tables based on Session 1 were computed. The differences were rounded to the nearest whole numbers, which were used as the uniform additive adjustment applied to the CBT students within each test.

For Grade 3 students, who do not have prior year scores, there was no propensity score matched samples being generated. Alternatively, the average of mode adjustments of Grades 4 and 5 was computed and used as the adjustment for Grade 3 CBT students. The CBT adjustments in all grade/subject were summarized in Table D.7.

Table D.7. CBT Scale Score Adjustments Comparison after Matching

Subject	Grade	Session	*PBT*			*CBT*			Δ	*d*	CBT Adjustment	
			n	***M***	***SD***	***n***	***M***	***SD***			**Based on Session 1 + 2**	**Based on Session 1**
ELA	3	1+2	-	-	-	-	-	-	-2.36	-	+2	+2
		1	-	-	-	-	-	-	-2.00	-		
	4	1+2	11539	599.63	18.64	11539	598.99	17.34	-0.64	-0.03	+1	+2
		1	11539	604.21	23.30	11539	602.55	22.07	-1.66	**-0.07**		
	5	1+2	25210	600.01	20.41	25210	597.76	19.85	-2.25	**-0.11**	+2	+2
		1	25210	601.60	22.45	25210	599.26	22.56	-2.34	**-0.10**		
	6	1+2	14647	601.35	18.57	14647	600.57	17.50	-0.78	-0.04	+1	+1
		1	14647	603.75	20.44	14647	602.44	19.60	-1.31	**-0.06**		
	7	1+2	11736	600.77	19.03	11736	599.82	17.19	-0.95	-0.05	+1	+2
		1	11736	602.54	21.22	11736	601.00	20.09	-1.54	**-0.07**		
	8	1+2	19875	598.94	20.01	19875	597.99	18.61	-0.95	-0.05	+1	+2
		1	19875	601.51	21.49	19875	599.76	20.97	-1.75	**-0.08**		
Math	3	1+2	-	-	-	-	-	-	-0.54	-	+1	+1
		1	-	-	-	-	-	-	-0.71	-		
	4	1+2	16520	599.74	19.31	16520	598.43	18.49	-1.31	**-0.07**	+1	+1
		1	16520	600.02	19.51	16520	599.39	19.10	-0.63	-0.03		
	5	1+2	17660	600.67	18.74	17660	598.98	18.53	-1.69	**-0.09**	+2	+1
		1	17660	600.95	18.94	17660	600.16	18.75	-0.79	-0.04		
	6	1+2	9840	604.07	17.45	9840	602.47	16.65	-1.60	**-0.08**	+2	+1
		1	9840	604.09	17.31	9840	602.74	16.56	-1.35	**-0.07**		
	7	1+2	18446	602.61	18.59	18446	600.38	18.32	-2.24	**-0.12**	+2	+2
		1	18446	602.22	19.91	18446	600.48	19.58	-1.74	**-0.08**		
	8	1+2	10133	599.27	19.22	10133	597.74	18.54	-1.53	**-0.07**	+2	+1
		1	10133	599.98	17.99	10133	599.26	17.03	-0.72	-0.04		

Note. The standardized difference (*d*) with an absolute value greater than 0.05 after matching was bolded.

D.4.2. Conclusions

Following the previous administrations, two administration modes were offered in the Grades 3–8 ELA and Mathematics tests in Spring 2021. The decision to offer PBT vs. CBT was optional, and the population of students who tested via CBT were not assumed equivalent to the population of students who tested via PBT. The same matched samples from the original-year mode comparability studies were used in this study. The results revealed small differences between CBT and PBT group across the shortened tests based on Session 1 only, similar to the results across the original tests including both Sessions 1 and 2. For both the original test and the shortened test, mode effects slightly favor PBT groups. The observed differences of the scale scores from the new RSSS tables based on Session 1 were applied as the adjustments to CBT students to ensure that students received comparable test scores regardless of the test mode.

Appendix E: Memo on Operational Test Mode Comparability

THE STATE EDUCATION DEPARTMENT / THE UNIVERSITY OF THE STATE OF NEW YORK / ALBANY, NY 12234

Assistant Commissioner
Office of State Assessment

August 2021

TO: District Superintendents
Superintendents of Schools
Principals of Public, Religious, and Independent Schools
Leaders of Charter Schools

FROM: Steven E. Katz

SUBJECT: Comparability of Spring 2021 Grades 3-8 English Language Arts and Mathematics Paper-based and Computer-based Tests

This memorandum provides information regarding the comparability of the Spring 2021 Grades 3–8 English Language Arts (ELA) and Mathematics paper-based and computer-based tests.

Background

Being mindful of the variability in the manner that instruction was being provided to students across the State in response to COVID-19, for the Spring 2021 Grades 3-8 English Language Arts and Mathematics Tests, schools were only required to administer Session 1 of the ELA and Mathematics Tests to students, and Session 1 tests were reduced in length by the removal of the embedded multiple-choice field test questions. NYSED also administered pre-existing test forms from one of the last two administrations (either 2018 or 2019) for the 2021 administration of Grades 3–8 ELA and Mathematics Tests.

The Department offered the assessments in two administration modes: paper-based testing (PBT) and computer-based testing (CBT). Administering these tests via CBT was optional for schools and those schools that chose to offer CBT made this decision independently for each subject and grade. As in previous years, the Department provided readiness verification tools to help those schools selecting CBT ensure they were well equipped and prepared to provide a successful CBT experience for their students. Additionally, several CBT practice test sessions were made available to CBT schools to familiarize students and teachers with the CBT delivery system. Each of the CBT practice test sessions featured examples of the types of test questions included on the tests. This provided the opportunity for students to practice answering ELA and mathematics test questions on the computer devices they would be using for the actual test.

To further ensure fairness, the Department's contractor, Questar Assessment Inc., addressed issues of comparability on the shortened 2021 tests (Session 1 only), to identify whether or not there were any differences in student performance that could be attributed to the mode of test administration (i.e., PBT versus CBT).

The methodology to address issues of comparability and results for the Spring 2021 Grades 3-8 ELA and Mathematics Tests are summarized below.

Comparability Methodology

In response to COVID-19, NYSED provided schools much more flexibility than had been necessary in past years in the local in-school scheduling and administration of the Spring 2021 ELA and Mathematics Tests. Given these unique circumstances, schools were not expected to bring students into the building to participate in the tests if the students were receiving entirely remote instruction during the testing windows. As a result, the population of students testing in 2021 was not representative of the student population overall. Therefore, issues of comparability were best addressed by the Department's contractor, Questar Assessment Inc., using data associated with the initial administration year for the assessments given (i.e., 2018 or 2019).

For the 2018 and 2019 administrations, a method called propensity score matching was employed to compare those students who tested via PBT with those who tested via CBT. Propensity score matching allowed for the identification of groups of students who tested via PBT that were similar to the groups of students who tested via CBT on a number of school and student characteristics, including achievement on the prior year's test. (Please see the "Comparability of Spring 2018 Grades 3-8 English Language Arts and Mathematics Paper-based and Computer-based Tests" or the "Comparability of Spring 2019 Grades 3-8 English Language Arts and Mathematics Paper-based and Computer-based Tests" for more information.)

Questar used these previously matched groups of PBT and CBT students to make a direct comparison of student results between the two groups for Session 1 only. For comparison, only the mean scale scores from Session 1 were calculated for each grade and subject by mode of testing. The results are shown in the section below.

Results

Table 1 shows the Session 1 scale score means for the PBT and CBT groups from either the 2018 or 2019 English Language Arts Tests by grade (depending on the form administered), as well as the differences in mean scale scores between the matched groups. Table 2 shows these same data for the Mathematics Tests.

For those tests in which a difference in mean scale scores between the two comparable groups was observed for Session 1, the scale scores for all students who took the test in that grade via CBT for the Spring 2021 administration were adjusted by adding the number of scale score points shown in the "Difference" columns of Tables 1 and 2 to the CBT students' scale scores, up to the maximum attainable scale score. Table 3 presents a summary of the adjustments applied for the Spring 2021 administration of the Grades 3-8 ELA and Mathematics Tests.

These slight adjustments ensured that students who demonstrated comparable proficiencies in their knowledge and skills received comparable scores whether they tested on paper or on computer in 2021.

For questions concerning the Grades 3-8 ELA or Mathematics Tests, please email the Office of State Assessment or call 518-474-5902. For questions concerning CBT, please email CBT Support.

Table 1. *PBT and CBT Means and Differences for Grades 3-8 ELA, Session 1 only*

	PBT Scale Score Mean	CBT Scale Score Mean	Difference (Rounded to nearest whole number)
Grade 3	See footnote*		n/a
Grade 4	604.2	602.60	+2
Grade 5	601.6	599.3	+2
Grade 6	603.8	602.4	+1
Grade 7	602.5	601.0	+2
Grade 8	601.5	599.8	+2

* Because Grade 3 students have no prior test results on which to match PBT to CBT students, a PBT comparison group was not created and group means were not calculated for this grade level.

Table 2. *PBT and CBT Means and Differences for Grades 3-8 Mathematics, Session 1 only*

	PBT Scale Score Mean	CBT Scale Score Mean	Difference (Rounded to nearest whole number)
Grade 3	See footnote*		n/a
Grade 4	600.0	599.4	+1
Grade 5	601.0	600.2	+1
Grade 6	604.1	602.7	+1
Grade 7	602.2	600.5	+2
Grade 8	600.0	599.3	+1

* Because Grade 3 students have no prior test results on which to match PBT to CBT students, a PBT comparison group was not created and group means were not calculated for this grade level.

Table 3. *Summary of Scale Score Adjustments for CBT*

	ELA Scale Score Adjustment	Math Scale Score Adjustment
Grade 3	+2*	+1*
Grade 4	+2	+1
Grade 5	+2	+1
Grade 6	+1	+1
Grade 7	+2	+2
Grade 8	+2	+1

* Because Grade 3 students have no prior test results on which to match PBT to CBT students, a PBT comparison group was not created and group means were not calculated for this grade level. Instead, the mean adjustment for the other elementary grades for which a comparison was possible (i.e., Grades 4 & 5) was applied to the scores of Grade 3 students who tested via CBT.