2020-2021 LEAP 2025 High School Operational Technical Report English I, English II, Algebra I, and Geometry

Submitted to the
Louisiana Department of Education

January 2022

DATA RECOGNITION
DRC
CORPORATION

This online-only document was published at a cost of $33,533. This document was published for the Louisiana Department of Education, P.O. Box 94064, Baton Rouge, LA 70804-9064, by Data Recognition Corporation, 13490 Bass Lake Road, Maple Grove, MN 55311. This material was printed in accordance with the standards for printing by State Agencies established pursuant to R.S. 43:31.

Table of Contents

Executive Summary

This report is a technical summary of the 2020-2021 administrations of the Louisiana Educational Assessment Program (LEAP 2025) in English language arts (ELA) and mathematics for high school. The LEAP 2025 summative assessments in ELA and mathematics are administered in grades 3 through 8 and high school. These tests are designed to measure students' readiness for the next grade or course of study and proficiency in ELA and mathematics. The ELA and mathematics test forms were developed by Data Recognition Corporation (DRC) test development staff using the New Meridian's item bank as well as items from the Louisiana Department of Education's own item bank. Items taken from these banks were on pre-established item response theory (IRT) scales. This section provides a summary of the 2020–2021 operational technical report.

E.1 Overview of This Report

This technical report documents the major activities of the testing cycle and provides details that confirm that the processes and procedures applied in the LEAP 2025 assessments adhered to appropriate professional standards and practices of educational assessment. Ultimately, this report serves to document evidence that valid inferences about Louisiana student performance in ELA and mathematics can be derived from the LEAP 2025 assessments. An overview of major activities documented within this report is provided below.

The Uses of Test Scores (Chapter 2)

Chapter 2 of the technical report discusses the concept of validity evidence. This technical report is composed of evidence that supports the intended uses of the LEAP 2025 test scores, and Chapter 2 discusses some of those uses.

Test Content Development (Chapter 3)

Chapter 3 of the technical report provides a summary of the test development activities that occurred to create the 2020–2021 operational test forms.

Test Administration (Chapter 4)

Chapter 4 of the technical report describes the processes implemented and the information disseminated to help ensure standardized test administration procedures and, thus, uniform test administration conditions for students.

Constructed-Response and Technology-Enhanced Scoring (Chapter 5)

Chapter 5 of the technical report describes the processes used to score constructed-response and technology-enhanced items. This chapter discusses how scorers are trained and the measures used to ensure consistency among scorers. Finally, this chapter presents the results of the inter-rater reliability studies.

Operational Data Analyses (Chapter 6)

Chapter 6 of the technical report includes a detailed description of the operational data analyses of the 2020–2021 LEAP 2025 ELA and mathematics assessments, which include the following major parts: the classical item analysis; calibration, scaling, and linking using IRT models; and student scoring. This chapter also describes the demographics of the calibration samples and compares them to state census data. It reports the results of the classical item analysis and the results of the calibration, scaling, and linking processes.

Test Results (Chapter 7)

Chapter 7 of the technical report contains information on the results of the 2020–2021 LEAP 2025 assessments. Detailed summary statistics based on scale scores and achievement levels are also provided. Finally, this chapter presents information on the score reports sent to school systems.

Performance-Level Setting (Chapter 8)

Chapter 8 of the technical report briefly discusses performance-level setting. It provides a brief overview of the procedures for performance-level setting and for derivation of the cut scores used to classify students into achievement levels for ELA and mathematics.

Evidence of Construct-Related Reliability (Chapter 9)

Chapter 9 of the technical report provides evidence of the reliability and validity of the LEAP 2025 test scores. This chapter provides detailed evidence of the reliability of the tests and information on the decision consistency of the cut scores. It also provides evidence of construct validity for the LEAP 2025 test scores.

Fairness (Chapter 10)

Chapter 10 of the technical report discusses fairness and how the LEAP 2025 assessments are constructed to be fair to all Louisiana students. This chapter summarizes the results of the differential item functioning (DIF) analysis. It also discusses the results of an impact analysis designed to determine whether large differences exist within the test results of different demographic groups in Louisiana.

E.2 Administration

Louisiana administered the LEAP 2025 summative assessments in ELA and mathematics to high school students in 2020–2021. The forms used for spring 2021 administration were previously administered in the spring of 2019. Computer-based tests (CBT) were administered during the following three testing windows: December 1 through 18, 2020, or January 6-26, 2021; April 15 through May 21, 2021; and June 21–25, 2021. Test administration is discussed in Chapter 4 of this report.

Ninety-seven school systems and twenty-nine charter schools administered the ELA and mathematics LEAP 2025 high school tests across the three administrations.

E.3 Student Performance

Table E.1 presents the percentage of students in 2020-2021 who were classified in each of the achievement levels for each subject. In general students that make up the population for each administration are:

- Fall: students from schools with block schedules and students retesting
- Spring: students from schools with block and regular schedules, as well as students retesting
- Summer: primarily students retesting

Table E.1 Percentage of Students Classified in Achievement Levels Using 2020-2021 Census Data: English Language Arts and Mathematics

Administration	Subject	Unsatisfactory	Approaching Basic	Basic	Mastery	Advanced
Fall	English I	21.6	22.8	23.1	25.5	7.1
	English II	20.5	19.6	22.1	28.5	9.2
	Algebra I	17.7	30.1	28.2	22.8	1.2
	Geometry	6.9	34.3	31.4	23.2	4.3
Spring	English I	14.3	20.2	25.9	32.3	7.3
	English II	18.1	16.3	21.1	29.0	15.5
	Algebra I	17.2	27.2	26.4	26.2	3.0
	Geometry	6.2	32.6	32.6	24.6	4.0
Summer	English I	58.0	30.6	9.1	2.2	0.2
	English II	62.6	26.2	8.3	2.5	0.4
	Algebra I	34.9	45.4	16.4	2.7	0.5
	Geometry	18.3	65.5	14.0	1.9	0.3

More information on student performance may be found in Chapter 7 of this report.

E.4 Validity and Test Scores

Most sections of this technical report are designed to provide validity evidence to support the use of the LEAP 2025 test scores. Chapter 2 discusses the uses of the LEAP 2025 test scores. Chapter 3 discusses the test development process used to create the LEAP 2025 tests, which is important to the content-related validity of the LEAP 2025 test scores. Chapter 4 presents information on test administration. Chapter 5 discusses the scoring process and the results of the inter-rater reliability studies. Chapter 6 presents the test scaling and linking procedures, student scoring methodology, and the results of other operational data analyses. Chapter 7 reviews the results of the 2020-2021 administrations and gives an overview of the score reports that were electronically delivered to the school systems for distribution to schools and parents. Chapter 8 highlights the procedures for performance-level setting implemented by Partnership for Assessment of Readiness for College and Careers (PARCC), which were used because PARCC's standards and achievement levels were used for the LEAP 2025. Chapter 9 discusses reliability and construct-related validity. Chapter 10 gives an overview of the statistical processes used to evaluate bias to ensure the fairness of the LEAP 2025 for all examinees.

Chapter 1: Introduction

The LEAP 2025 assessment system is designed to measure students' knowledge of ELA, mathematics, science, and social studies. This report provides a technical overview of the LEAP 2025 ELA and mathematics high school assessments administered in the 2020-2021 academic year and presents evidence for the validity of the 2020-2021 LEAP 2025 ELA and mathematics high school assessment scores.

This chapter describes the background, purpose, and design of the LEAP 2025 assessments.

1.1 Background

In 2010, the Board of Elementary and Secondary Education (BESE) approved the Common Core State Standards (CCSS) in ELA and mathematics. After adopting the CCSS, Louisiana became a governing member of PARCC, a group of states working to develop high-quality assessments that measure the full range of the CCSS. Beginning in 2015, students in grades 3–8 began taking these newly aligned assessments.

In 2016, Louisiana ELA and mathematics academic content standards underwent a review process resulting in the adoption of the Louisiana Student Standards in English language arts and mathematics. In spring 2017, ELA and mathematics students in grades 3–8, except those qualifying for the LEAP Alternate Assessment Level 1 (LAA 1), took the LEAP 2025 assessments.

Beginning in the 2017–2018 school year, the Louisiana Department of Education (LDOE) transitioned to LEAP 2025 ELA and mathematics high school assessments, which were aligned to the Louisiana Student Standards in ELA and mathematics. The five-performance-level LEAP 2025 high school assessments replaced the four-performance-level End-of-Course (EOC) tests. Students enrolled in English I, English II, Algebra I, and Geometry took the LEAP 2025 high school assessments.

The information that follows describes the technical aspects of the 2020-2021 LEAP 2025 ELA and mathematics assessments and provides information about how to read and interpret the data on the 2020-2021 assessment reports.

1.2 Purpose of the LEAP 2025

The BESE and the LDOE are committed to ensuring that every student is on track to be successful in either postsecondary education or the workforce through their comprehensive plan Believe to Achieve. The LEAP 2025 supports this vision by measuring the full range of student performance and providing information for educators and parents about student readiness for college and careers.

1.3 Design of the LEAP 2025

High school students were administered computer-based tests (CBTs) in both ELA and mathematics. Additionally, a braille form was available for each course and content area. Online tools allowed students to magnify assessment items as needed. See Section 3.5 in Chapter 3 for more information about the accommodations and designated supports available for students taking the LEAP 2025. All mathematics assessments were translated into Spanish forms.

The 2020-2021 LEAP 2025 test blueprints and test design are based on the ELA https://resources.newmeridiancorp.org/ela-test-design/ and mathematics https://resources.newmeridiancorp.org/math-test-design/ blueprints of New Meridian's full forms. The 2020-2021 LEAP 2025 test blueprints and test design for ELA and mathematics differ from the New Meridian

blueprints and design in order to reduce testing time while maintaining full coverage and including a variety of standards.

The 2020-2021 LEAP 2025 ELA blueprints kept a similar design as the design of New Meridian's full form, including both performance-based tasks and stand-alone passage sets; however, only two of the three types of performance tasks—Research Simulation Task and Literary Analysis Task or Narrative Writing Task—are included on each of the LEAP 2025 English I and English II assessments. All three task types are represented across administrations, which encourages teachers to focus equally on all three writing types. Besides having two (instead of three) performance tasks, the 2020-2021 LEAP 2025 ELA blueprints are also different with respect to testing time and percentage of reading and writing points. Since the choice of Literary Analysis Task or Narrative Writing Task is determined during the forms construction process, alternative blueprints—one with a Literary Analysis Task and a Research Simulation Task and the other with a Research Simulation Task and a Narrative Writing Task—were created for each course's assessment.

The passages chosen for the 2020-2021 LEAP 2025 English I and English II assessments contain a variety of texts of different genres with a balance of authors by gender and ethnicity. The assessments also contain texts that appeal to a diverse student population. Chosen passages are authentic and contain a variety of different types of text with varying degrees of text complexity. They are rich in content, engaging, high-quality, and challenging. Additionally, paired passages are selected with careful consideration of the standards that require the use of more than one text. This combination of criteria during passage selection allows students to demonstrate their ability to read and comprehend a range of complex texts and helps ensure as much coverage of the standards as possible.

The LEAP 2025 ELA assessments focus on an integrated approach to reading and writing that reflects instruction in an effective ELA classroom and measures students' ability to understand what they read and express that understanding in writing. This means careful, close reading of complex grade-level literary and informational texts; a full range of texts from across the disciplines, including science, social studies, and the arts; tasks that integrate key ELA skills by asking students to read texts, answer reading and vocabulary questions about the texts, and then write using evidence from what they have read; questions worth answering, ordered in a way that builds meaning; a focus on students citing evidence from texts when answering questions about a specific passage or when writing about a set of related passages; and a focus on words that matter most in texts, that are essential to understanding a particular text, and that include context that allows students to determine literal and figurative meanings.

The 2020-2021 LEAP 2025 mathematics blueprints kept a similar design as the design of New Meridian's full form, with a few notable exceptions:

- Both assessment designs have three sessions, with Session 1 split into non-calculator and calculator sections. However, New Meridian's full form has three sessions that last 90 minutes each (for a total of 270 minutes), while LEAP 2025 has three sessions that last 80 minutes each (for a total of 240 minutes).
- In Algebra I, both assessment designs have the same number of Type II items worth 4 points. The LEAP 2025 design uses 1 more Type I item worth 1 point, 2 fewer Type I items worth 2 points, 1 fewer Type I item worth 4 points, 1 fewer Type II item worth 3 points, 1 more Type III item worth 3 points, and 1 fewer Type III item worth 6 points.
- In Geometry, both assessment designs have the same number of Type II items worth 4 points. The LEAP 2025 design uses 1 fewer Type I item worth 1 point, 1 fewer Type I item worth 2 points, 1 fewer Type I item worth 4 points, 1 fewer Type II item worth 3 points, 1 more Type III item worth 3 points, and 1 fewer Type III item worth 6 points.

- The LEAP 2025 mathematics assessments focus on testing the Louisiana Student Standards for Mathematics (LSSM) according to the components of rigor reflected in high-quality mathematics instructional tasks that
  - require students to demonstrate understanding of mathematical reasoning in mathematical and applied contexts;
  - assess accurate, efficient, and flexible application of procedures and algorithms;
  - rely on application of procedural skill and fluency to solve complex problems; and
  - require students to demonstrate mathematical reasoning and modeling in real-world contexts.

The LSSM support students in becoming mathematically proficient by focusing on three components of rigor: conceptual understanding, procedural skill and fluency, and application.

- Conceptual understanding refers to understanding mathematical concepts, operations, and relations. It is more than knowing isolated facts and methods. Students should be able to make sense of why a mathematical idea is important and the kinds of contexts in which it is useful. It also allows students to connect prior knowledge to new ideas and concepts.
- Procedural skill and fluency is the ability to apply procedures accurately, efficiently, and flexibly. It requires speed and accuracy in calculation while giving students opportunities to practice basic skills. Students' ability to solve more complex application tasks is dependent on procedural skill and fluency.
- Application provides a valuable context for learning and solving problems in a relevant and meaningful way. It is through real-world application that students learn to select an efficient method to find a solution, determine whether the solution(s) makes sense by reasoning, and develop critical thinking skills.

Each item on the LEAP 2025 Algebra I and Geometry assessments is referred to as a task and is identified by one of three types: Type I, Type II, or Type III. The tasks on the LEAP 2025 mathematics tests are aligned directly to the LSSM for all reporting categories.

- Type I tasks, designed to assess conceptual understanding, fluency, and application, are aligned to the major, additional, and supporting content for each grade. Some Type I tasks may be further aligned to LEAP 2025 evidence statements for the Major Content and Additional & Supporting reporting categories and allow for the testing of more than one LSSM on a single task.
- Type II tasks are designed to assess student reasoning ability of selected major content for the grade or the previous grade in applied contexts. Type II tasks are further aligned to LEAP 2025 evidence statements for the Expressing Mathematical Reasoning and Modeling & Application reporting categories.
- Type III tasks are designed to assess student modeling ability of selected content for the grade or the previous grade in applied contexts. Type III tasks are further aligned to LEAP 2025 evidence statements for the Expressing Mathematical Reasoning and Modeling & Application reporting categories.

Each of the three task types is aligned to one of four reporting categories: Major Content, Additional & Supporting Content, Expressing Mathematical Reasoning, or Modeling & Application. Each task type is designed to align with at least one of the Louisiana Student Standards for Mathematical Practice (MP). Additional details about the design of the ELA and mathematics assessments can be found in Chapter 3.

Chapter 2: The Uses of Test Scores

Validity is the central component of the LEAP 2025 assessments. The following excerpt is from the *Standards for Educational and Psychological Testing* (American Educational Research Association [AERA], American Psychological Association [APA], & National Council on Measurement in Education [NCME], 2014):

Ultimately, the validity of an intended interpretation of test scores relies on all the available evidence relevant to the technical quality of a testing system. Different components of validity evidence include evidence of careful test construction; adequate score reliability; appropriate test administration and scoring; accurate score scaling, equating, and standard setting; and careful attention to fairness for all test takers, as appropriate to the test interpretation in question (22).

As stated by the *Standards*, the validity of a testing program hinges on the use of the test scores. Validity evidence that supports the use of the LEAP 2025 test scores is provided in this technical report. This chapter examines some possible uses of the LEAP 2025 test scores. However, this technical report cannot anticipate all possible interpretations and uses of the LEAP 2025 test scores.

2.1 Uses of Test Scores

To understand whether a test score is being used properly, one must understand the purpose of the test. The intended uses of the LEAP 2025 test scores include the following:

- evaluating students' overall proficiency of the Louisiana Student Standards
- identifying students' strengths and weaknesses
- evaluating programs at the school, school system, and/or state level
- informing stakeholders, including students, teachers, school administrators, school system administrators, LDOE staff members, parents, and the public, of the status of students' progress toward meeting college- and career- readiness standards

This technical report refers to the uses of test-level scores (i.e., scale scores and achievement levels), category-level scores and achievement-level classifications, and subcategory-level scores and achievement-level classifications.

2.2 Test-Level Scores

At the test level, an overall scale score that is based on student performance on the entire test is reported. In addition, an associated level of achievement is reported. These scores and achievement levels indicate, in varying ways, a student's achievement. Test-level scores are reported at four reporting levels: the state, the school system, the school, and the student.

The LEAP 2025 high school ELA and mathematics test forms were developed by DRC's test development staff using New Meridian's item bank as well as items from the Louisiana Department of Education's own item bank. Items taken from these banks were on pre-established item response theory (IRT) scales for ELA and mathematics and were reviewed and approved for use by LDOE content experts and committees of Louisiana educators. Braille forms and Spanish translations of mathematics forms were also developed. See Chapter 3, "Test Content Development," for additional details about the processes used to develop these test forms.

The following sections discuss two types of test-level scores that are reported to indicate a student's achievement on the LEAP 2025 assessments: the scale score and its associated level of achievement.

2.3 Scale Scores

A scale score indicates a student's total performance on the LEAP 2025 assessments. The overall scale score quantifies the achievement being measured by the assessments. In other words, the scale score represents the student's level of achievement, where higher scale scores indicate higher levels of achievement on the test and lower scale scores indicate lower levels of achievement. For all LEAP 2025 test forms, the lowest obtainable scale score (LOSS) is 650 and the highest obtainable scale score (HOSS) is 850.

Scale scores are derived from raw scores (i.e., the number of items answered correctly). Raw scores depend on the items in a particular form of a test and can only be interpreted in terms of that particular set of test questions. This does not allow year-to-year or form-to-form comparison. Scale scores are more meaningful than raw scores because they maintain their meaning year-to-year, thus allowing comparisons of different test forms across the entire range of the ability scale.

2.4 Levels of Achievement

A student's performance on the LEAP 2025 assessments is reported in one of five levels of achievement: *Advanced, Mastery, Basic, Approaching Basic*, or *Unsatisfactory*. The cut scores for the ELA and mathematics achievement levels were established by PARCC using the Evidence-Based Standard Setting (EBSS) method (Beimers, Way, McClarty, & Miles, 2012) for the PARCC Performance-Level Setting (PLS) process. Details regarding the PLS process can be found in the *Performance Level Setting Technical Report* (Pearson, 2015).

Descriptions of each level of achievement, in terms of what a student should know and be able to do, are provided with the *LEAP 2025 Interpretive Guide* (see Chapter 7).

2.5 Use of Test-Level Scores

The LEAP 2025 scale scores and achievement levels provide summary evidence of student performance relative to the Louisiana Student Standards. Classroom teachers may use these scores as evidence of student achievement in English I, English II, Algebra I, and Geometry. At the aggregate level, school system and school administrators may use this information for activities such as curriculum planning. The results presented in this technical report provide evidence that the scale scores and achievement levels are valid and reliable indicators of what students know, understand, and are able to do relative to the Louisiana Student Standards in ELA and mathematics.

2.6 Category- and Subcategory-Level Subscores

A student's performance on the ELA reporting categories (i.e., reading and writing) is reported by one of three ratings: *Strong*, *Moderate*, or *Weak*.

Additionally, subcategory subscores are reported at the student level for ELA and mathematics. ELA has three subcategories for reading and two subcategories for writing, as described in Table 3.1, ELA Reporting Categories and Subcategories. Mathematics has four subcategories, as described in Table 3.6, Overview of LEAP 2025 Mathematics Task Types and Reporting Categories. Subcategory performance is reported in one of three ratings: *Strong*, *Moderate*, or *Weak*.

Although the performance ratings are determined only by the items included within a category or subcategory, the level of knowledge and ability needed to demonstrate a performance rating is connected to the level of knowledge and ability required by the assessments: a *Strong* rating requires similar knowledge and ability as the Mastery or Advanced achievement levels, a *Moderate* rating requires similar knowledge and ability as the Basic achievement level, and a *Weak* rating requires similar knowledge and ability as the Unsatisfactory and Approaching Basic achievement levels.

2.7 Use of the Category- and Subcategory-Level Ratings

The purpose of reporting category- or subcategory-level subscores on LEAP 2025 assessments is to show, for each student, the relationship between the overall achievement being measured and the skills in each of the areas defined by the reporting categories and subcategories. Teachers may use these ratings for individual students as indicators of strengths and weaknesses, but they are best corroborated by other evidence, such as grades, teacher feedback, and scores on other tests. Chapter 3 of this technical report provides evidence of content validity that supports the use of the category- or subcategory-level subscores. Chapter 9 of this technical report provides evidence of construct-related validity that further supports the use of these subscores.

Chapter 3: Test Content Development

Content-related validity in achievement tests is evidenced by a correspondence between test content and the range of knowledge and skills that compose the construct the assessment is designed to measure (i.e., the ELA or mathematics Louisiana Student Standards). Content-related validity can be demonstrated through consistent adherence to test blueprints, through a high-quality test development process that includes review of items for accessibility to English Learners and students with disabilities, and through alignment studies performed by independent groups. This chapter provides a detailed discussion of the test development process. In particular, it shows how rigorous procedures were followed to construct tests that reflect the full range of content that the 2020-2021 LEAP 2025 high school assessments were expected to cover.

This chapter is particularly relevant to the following sections of the *Standards for Educational and Psychological Testing* (American Educational Research Association [AERA], American Psychological Association [APA], & National Council on Measurement in Education [NCME], 2014): Standards 4.0, 4.1, and 4.7. It also addresses Standards 3.1, 3.2, 3.9, and 4.12, which are discussed in pertinent sections of this chapter.

Standard 4.0 states the following:

> Tests and testing programs should be designed and developed in a way that supports the validity of interpretations of the test scores for their intended uses. Test developers and publishers should document steps taken during the design and development process to provide evidence of fairness, reliability, and validity for intended uses for individuals in the intended examinee population (85).

Standard 4.1 states the following:

> Test specifications should describe the purpose(s) of the test, the definition of the construct or domain measured, the intended examinee population, and interpretations for intended uses. The specifications should include a rationale supporting the interpretations and uses of test results for the intended purpose(s) (85).

The 2020-2021 LEAP 2025 high school test specifications consisted of a blueprint and a design for each of the following tests: English I, English II, Algebra I, and Geometry. The 2020-2021 blueprints and test designs were closely aligned to blueprints of New Meridian's full forms. The test blueprints for the 2020-2021 LEAP 2025 high school ELA assessments were designed with the goal of having all students read, understand, and express their understanding of complex, grade-level texts. The test blueprints for the 2020-2021 LEAP 2025 mathematics assessments were designed with the goal of supporting students to become mathematically proficient by focusing on three components of rigor: conceptual understanding, procedural skill and fluency, and application. The 2020-2021 LEAP 2025 high school ELA and mathematics assessments provided questions that were reviewed by Louisiana educators to ensure their alignment to the Louisiana Student Standards and appropriateness for Louisiana students, measured the full range of student performance, and informed educators and parents about student readiness in ELA and mathematics and whether students are "on track" for college and careers. For ELA and mathematics, the 2020-2021 LEAP 2025 assessments use the same blueprints and reporting categories and subcategories that were used in 2018-2019.

To construct the assessments following the LDOE-approved test blueprints and test designs, LDOE and DRC collaborated to use items from the New Meridian and Louisiana-owned item banks. Both item banks are

comprised of items aligned to the Louisiana Student Standards. DRC contracted with New Meridian and was provided access to the entire bank of items and passage sets that could potentially be used on operational forms. The acquired items and passages and the Louisiana-owned items and passage sets made up the available item pool used for the 2020-2021 LEAP 2025 high school forms construction. LDOE and DRC confirmed that all items selected for use on the LEAP 2025 forms were appropriate for use on Louisiana assessments by convening committees of Louisiana educators who reviewed and approved items from the item banks prior to form selection.

The ELA and mathematics LEAP 2025 assessments for high school were developed based on the requirements of "RFP #678PUR-LEAP 2025 English Language Arts and Mathematics Assessment System" as follows:

The assessments shall be

- aligned to the ELA and mathematics Louisiana Student Standards;
- designed to be accessible for use by the widest possible range of students, including, but not limited to, students with disabilities and students with limited English proficiency [English Learners];
- constructed to yield valid and reliable test results;
- constructed to report student performance using achievement level policy definitions and reporting categories that are comparable to a significant number of other states;
- developed to limit the amount of testing time required and to be in compliance with all state laws regarding testing time;
- developed and reviewed with Louisiana educator involvement;
- non-computer adaptive;
- used in assessing students' readiness to successfully transition to postsecondary education and the workplace; and
- administered, scored, and reported through a separate administration contract.

The products of the above requirements are computer-based tests (CBTs) comprised of New Meridian and Louisiana-owned items aligned to the Louisiana Student Standards. Louisiana had access to the complete New Meridian item bank for forms administered during the 2020-2021 school year. Items and passage sets were deemed appropriate for use on Louisiana assessments by Louisiana educators during an item alignment review. These items and passage sets were approved because they aligned to the Louisiana Student Standards and/or Louisiana Evidence Statements for mathematics and because they were free of issues related to bias, fairness, and sensitivity. These items and passage sets became the available item pool used to construct the forms administered during the 2020-2021 school year. For each course, forms administered were selected from the available pool of items and/or passage sets. DRC and LDOE content experts scrutinized each final blueprint to ensure optimal content coverage and prudent use of time and resources. In general, the blueprints represent content sampling proportions that reflect intended emphasis in instruction and mastery in each course and are comparable to the New Meridian full form test blueprints. The test specifications provide the numbers of items by reporting category, assessment focus, and item type, and they demonstrate the desired proportions within test delivery and available item pool constraints. These specifications can be found in the *2020-2021 LEAP 2025 High School English Language Arts and Mathematics Assessment Frameworks*. All assessments were fixed forms, which means that all students who received the same form were administered the same set of items, as the forms were not adaptive.

The LEAP 2025 high school assessments are administered in fall, spring, and summer each school year. For fall and summer administrations, two forms are administered: an operational form and an administrative

error form, which is used only if there is an administrative testing error (see Chapter 4 for additional details regarding the administrative error form). For spring administrations, one operational form and one administrative error form are administered. Typically, a senior form is also administered, but due to using an intact form in 2021, the senior form was not necessary. The forms are administered on a rotating schedule, so they are not the same from administration to administration.

3.1 Defining the Specific Test Blueprint

The test blueprints for the 2020-2021 assessments were designed based on two primary factors: (1) the content requirements of the Louisiana Student Standards and (2) the reporting needs of the assessments.

English I and English II Test Blueprints and Test Designs

The English I and English II tests were administered during operational testing windows: December 1 through 18, 2020, or January 6-26, 2021; April 15 through May 21, 2021; and June 21–25, 2021. Only two of the three types of performance tasks—Research Simulation Task, Literary Analysis Task, and Narrative Writing Task—were included on each of the Louisiana tests. All three types were represented across administrations (fall, spring, and summer), which encourages educators to focus on all three writing types. Since the choice of Literary Analysis Task or Narrative Writing Task is determined during the forms construction process, alternative blueprints—one with a Literary Analysis Task and a Research Simulation Task and the other with a Research Simulation Task and a Narrative Writing Task—are created for each course.

Student performance on the LEAP 2025 high school ELA assessments is reported by category and subcategory as outlined in the following table.

Table 3.1 ELA Categories and Subcategories

Category	Subcategory	Subcategory Description
Reading	Reading Literary Text	Students read and demonstrate comprehension of grade-level fiction, drama, and poetry.
	Reading Informational Text	Students read and demonstrate comprehension of grade-level nonfiction, including texts about history, science, art, and music.
	Reading Vocabulary	Students use context to determine the meaning of words and phrases in grade-level texts.
Writing	Written Expression	Students use details from provided texts to compose well-developed, organized, clear writing.
	Knowledge and Use of Language Conventions	Students use the rules of standard English (grammar, mechanics, and usage) to compose writing.

These reporting categories provide parents and educators with valuable information about

- overall student performance, including readiness to continue further study in English language arts;
- student performance broken down by subcategory, which may help identify when students need additional support or more challenging work in reading and writing; and
- how well schools and school systems help students achieve expectations.

The session testing times shown in the ELA test designs (see Tables 3.2 and 3.3) are based on New Meridian testing times proportioned to be comparable based on the passage type being tested. The passage set that comes after the Narrative Writing Task or the Literary Analysis Task is designed to balance the reading load between the Narrative Writing Task or the Literary Analysis Task and the Research Simulation Task and to provide consistent timing in sessions 1 and 2.

Table 3.2 English I and English II Test Design—Literary Analysis Task and Research Simulation Task

Session	Task/ Item Set	Number of Passages	Categories/ Subcategories	Number of Two-Point SR Items	Number of Points from Two-Point SR Items	Number of PCR Items	Number of Points from PCR Items	Total Number of Items	Total Number of Points	Assessable ELA Student Standards (by subcategory)	Testing Time (minutes)
1	Literary Analysis Task	2	Reading: Reading Literary Text/Reading Vocabulary*	6	12	1	4	6	16	RL Standards 1-3, 5-10; vocabulary standards RL.4, L.4, L.5	90
			Writing: Written Expression	0	0		12	1	12	Writing standards W.1-2, 4, 9, 10	
			Writing: Knowledge and Use of Language Conventions	0	0		3		3	Convention standards L.1, 2, plus language skills from previous grades	
	Reading Literary / Informational Texts	1	Reading (Reading Literary Text/Reading Informational Text/Reading Vocabulary*)	4	8	0	0	4	8	RL Standards 1-3, 5-10; RI standards 1-3, 5-10; vocabulary standards RL.4, RI.4, L.4, L.5	
	Totals	3		10	20	1	19	11	39		
2	Research Simulation Task	3	Reading: Reading Informational Text/ Reading Vocabulary*	8	16	1	4	8	20	RI standards 1-3, 5-10; vocabulary standards RI.4, L.4, L.5	90
			Writing: Written Expression	0	0		12	1	12	Writing standards W.1-2, 4, 7- 10,	
			Writing: Knowledge and Use of Language Conventions	0	0		3		3	Convention standards L.1, 2, plus language skills from previous grades	
	Totals	3		8	16	1	19	9	35		
3	Reading Literary Texts	2-3	Reading: Reading Literary Text/Reading Vocabulary*	10	20	0	0	10	20	RL Standards 1-3, 5-10; vocabulary standards RL.4, L.4, L.5	80**
	Reading Informational Texts		Reading: Reading Informational Text/Reading Vocab*			0	0			RI standards 1-3, 5-10; vocabulary standards RI.4, L.4, L.5	
	Totals	2-3		10	20	0	0	10	20		
English I & II Totals		8-9	Reading: Reading Literary Text/Reading Vocabulary*	28	56	2	4	28	64	64	260
			Reading: Reading Informational Text/Reading Vocabulary*				4				
			Writing: Written Expression	0	0		24	2	24	30	
			Writing: Knowledge and Use of Language Conventions	0	0		6		6		
			Total	28	56	2	38	30	94	94	

*Reading vocabulary items must constitute at least eight points on the test.
**The time in session 3 allows for an additional passage set that is a field-test or placeholder passage set.

Table 3.3 English I and English II Test Design—Research Simulation Task and Narrative Writing Task

Session	Task/ Item Set	Number of Passages	Categories/ Subcategories	Number of Two-Point SR Items	Number of Points from Two-Point SR Items	Number of PCR Items	Number of Points from PCR Items	Total Number of Items	Total Number of Points	Assessable ELA Student Standards (by subcategory)	Testing Time (minutes)
1	Research Simulation Task	3	Reading: Reading Informational Text/Reading Vocabulary*	8	16	1	4	8	20	RI standards 1-3, 5-10; vocabulary standards RI.4, L.4, L.5	90
			Writing: Written Expression	0	0		12	1	12	Writing standards W.1-2, 4, 7-10	
			Writing: Knowledge and Use of Language Conventions	0	0		3		3	Convention standards L.1, 2, plus language skills from previous grades	
	Totals	3		8	16	1	19	9	35		
2	Narrative Writing Task	1	Reading: Reading Literary Text/Reading Vocabulary*	4	8	1	0	4	8	RL Standards 1-3, 5-10; vocabulary standards RL.4, L.4, L.5	90
			Writing: Written Expression	0	0		12	1	12	Writing standards W.3, 4, 10	
			Writing: Knowledge and Use of Language Conventions	0	0		3		3	Convention standards L.1, 2, plus language skills from previous grades	
	Reading Literary / Informational Texts	1-2	Reading (Reading Literary Text/Reading Informational Text/Reading Vocabulary*)	6	12	0	0	6	12	RL Standards 1-3, 5-10; RI standards 1-3, 5-10; vocabulary standards RL.4, RI.4, L.4, L.5	
	Totals	2-3		10	20	1	15	11	35		
3	Reading Literary Texts	2-3	Reading: Reading Literary Text/Reading Vocabulary*	10	20	0	0	10	20	RL Standards 1-3, 5-10; vocabulary standards RL.4, L.4, L.5	80**
	Reading Informational Texts		Reading: Reading Informational Text/Reading Vocabulary*			0	0			RI.1-3, 5-10; vocabulary standards RI.4, L.4, L.5	
	Totals	2-3		10	20	0	0	10	20		
English I & II Totals		**7-9**	Reading: Reading Literary Text/Reading Vocabulary*	28	56	2	0	28	60	60	260
			Reading: Reading Informational Text/Reading Vocabulary*				4				
			Writing: Written Expression	0	0		24	2	24	30	
			Writing: Knowledge and Use of Language Conventions	0	0		6		6		
			Total	28	56	2	34	30	90	90	

*Reading vocabulary items must constitute at least eight points on the test.
**The time in session 3 allows for an additional passage set that is a field-test or placeholder passage set.

The LEAP 2025 high school ELA assessments consist of tasks and reading passage sets. The tasks are described below.

- Narrative Writing Task
    - This task asks students to read a literary text, answer a set of selected-response questions about the text, and create a narrative related to the text (e.g., finish the story or retell the story in another narrative form, such as a journal entry).
    - This task focuses on students' ability to use narrative elements (e.g., dialogue, description) when writing.
- Literary Analysis Task
    - This task provides students with an opportunity to show their understanding of literature. It asks students to read two literary texts, answer a set of selected-response questions about the texts, and write an extended response that compares and/or explains key ideas or elements in the texts (e.g., central idea/message, contribution of illustrations, characterization).
    - This task focuses on students' ability to read complex text closely and asks them to carefully consider literature worthy of close study.
- Research Simulation Task
    - This task mirrors the research process by presenting three texts on a given topic. Students answer a set of selected-response questions about the texts and then write an extended response about some aspect of the related texts (e.g., relationship between a series of events, ideas, or concepts; comparison/contrast of key details; presentation of information).
    - This task requires students to synthesize information from related informational resources.

The following item types were included in the 2020-2021 LEAP 2025 high school ELA assessments:

- Selected-Response Items:
    - Evidence-based selected response (EBSR): This item type consists of two parts. One part asks students to show their understanding of a text, and the other part asks students to identify evidence to support that understanding. The evidence supports a generalization, conclusion, or inference. This type of item is designed to provide students with opportunities to make explicit the evidence that supports their close analysis of a specific text.
    - Multiple select (MS): This item type requires students to select more than one correct answer and may appear as a one-part question or as part of an EBSR item. This type of item allows for the assessment of students' ability to identify multiple pieces of evidence to support a claim.
    - Technology enhanced (TE): This item type allows measurement of learning that may not be sufficiently measured by traditional multiple-choice items. TE items can measure the ordering of ideas within a summary; ordering of steps in a process; sorting, classifying, and categorizing ideas; matching of two themes/ideas to their unique evidence, etc. The technology used in TE items offers students additional ways to show understanding that parallels the classroom instructional techniques that teachers use to determine whether students are able to comprehend complex, grade-level text. TE Items may involve any of the following:

    - Highlighting text: requires students to select text-based answer(s) from within a larger text
    - Drag and drop: requires students to move draggable elements (e.g., words, phrases, or sentences) into one or more drop boxes (e.g., cells within a table or part[s] of a diagram)
    - Drop-down menu: requires students to select from one or more drop-down menus to complete a phrase or sentence
    - Match interaction table: requires students to select a checkbox in each row from two or more columns to classify statements presented in each row
- Prose constructed response (PCR): This item type appears at the end of each task and asks students to create an extended, complete written response. It elicits evidence that students have understood a text or texts they have read and can communicate that understanding well, both in terms of written expression and in terms of knowledge and use of language conventions.

A variety of item types allows for the measurement of the full range of student performance. Items and tasks should be clearly aligned to specific standards. Some items and tasks may ask students to draw evidence from one specific standard, while others may ask students to draw evidence from several standards.

The following tables detail the number of items and points by session and item type for English I and English II forms.

Table 3.4 Distribution of English I Items and Points by Session and Item Type

Form	Session	EBSR		MS		TE		PCR		Total No. of Pts.
		No. of Items	No. of Pts.	No. of Items	No. of Pts.	No. of Items	No. of Pts.	No. of Items	No. of Pts.	
A	1. Research Simulation Task	6	12	0	0	2	4	1	19	90
	2. Narrative Writing Task/ Reading Passage	5	10	2	4	3	6	1	15	
	3. Reading Literary/ Informational Texts	8	16	1	2	1	2	0	0	
B	1. Literary Analysis Task/ Reading Passage	8	16	0	0	2	4	1	19	94
	2. Research Simulation Task	6	12	0	0	2	4	1	19	
	3. Reading Literary/ Informational Texts	6	12	1	2	3	6	0	0	
C	1. Research Simulation Task	6	12	1	2	1	2	1	19	90
	2. Narrative Writing Task/ Reading Passage	5	10	3	6	2	4	1	15	
	3. Reading Literary/ Informational Texts	6	12	3	6	1	2	0	0	
D	1. Literary Analysis Task/ Reading Passage	9	18	0	0	1	2	1	19	94
	2. Research Simulation Task	4	8	2	4	2	4	1	19	
	3. Reading Literary/ Informational Texts	8	16	0	0	2	4	0	0	
E	1. Research Simulation Task	4	8	2	4	2	4	1	19	90
	2. Narrative Writing Task/ Reading Passage	4	8	2	4	4	8	1	15	
	3. Reading Literary/ Informational Texts	6	12	1	2	3	6	0	0	

Table 3.5 Distribution of English II Items and Points by Session and Item Type

Form	Session	EBSR		MS		TE		PCR		Total No. of Pts.
		No. of Items	No. of Pts.	No. of Items	No. of Pts.	No. of Items	No. of Pts.	No. of Items	No. of Pts.	
A	1. Research Simulation Task	5	10	2	4	1	2	1	19	90
	2. Narrative Writing Task/ Reading Passage	2	4	7	14	1	2	1	15	
	3. Reading Literary/ Informational Texts	5	10	4	8	1	2	0	0	
B	1. Literary Analysis Task/ Reading Passage	4	8	3	6	3	6	1	19	94
	2. Research Simulation Task	5	10	2	4	1	2	1	19	
	3. Reading Literary/ Informational Texts	4	8	4	8	2	4	0	0	
C	1. Literary Analysis Task/ Reading Passage	8	16	0	0	2	4	1	19	94
	2. Research Simulation Task	5	10	1	2	2	4	1	19	
	3. Reading Literary/ Informational Texts	2	4	4	8	4	8	0	0	
D	1. Literary Analysis Task/ Reading Passage	4	8	3	6	3	6	1	19	94
	2. Research Simulation Task	4	8	2	4	2	4	1	19	
	3. Reading Literary/ Informational Texts	6	12	1	2	3	6	0	0	
E	1. Research Simulation Task	4	8	2	4	2	4	1	19	90
	2. Narrative Writing Task/ Reading Passage	7	14	1	2	2	4	1	15	
	3. Reading Literary/ Informational Texts	2	4	5	10	3	6	0	0	

Mathematics Test Blueprints and Test Designs

The mathematics assessments were administered during operational testing windows: December 1 through 18, 2020 or January 6-26, 2021; April 15 through May 21, 2021; and June 21–25, 2021. The 2020-2021 mathematics assessments included three test sessions, and each test session included the four mathematics subcategories and the three mathematics task types. See Table 3.6 for details about categories and task types.

Each item on the LEAP 2025 mathematics assessment is referred to as a task and is identified by one of three types: Type I, Type II, and Type III. As shown in the following table, each task type is aligned to one or two of four reporting categories: Major Content, Additional & Supporting Content, Expressing Mathematical Reasoning, or Modeling & Application. Each task type is designed to align to at least one of the Standards for Mathematical Practice (MP).

Table 3.6 Overview of LEAP 2025 Mathematics Task Types and Reporting Categories

Task Type	Description	Reporting Categories	Mathematical Practice(s)
Type I	Conceptual understanding, fluency, and application	*Major Content:* solve problems involving the major content for the grade level. *Additional & Supporting Content:* solve problems involving the additional and supporting content for the grade level.	Can involve any or all practices
Type II	Written arguments/ justifications, critique of reasoning, or precision in mathematical statements	*Expressing Mathematical Reasoning:* express mathematical reasoning by constructing mathematical arguments and critiques.	Primarily MP.3 and MP.6 but may also involve any of the other practices
Type III	Modeling/application in a real-world context or scenario	*Modeling & Application:* solve real-world problems engaging particularly in the modeling practice.	Primarily MP.4 but may also involve any of the other practices

These reporting categories provide parents and educators with valuable information about

- overall student performance, including readiness to continue further study in mathematics;
- student performance broken down by mathematics subcategories, which may help identify when students need additional support or more challenging work; and
- how well schools and school systems help students achieve higher expectations.

Tables 3.7 and 3.8 provide the distribution of operational points by reporting category and by form for each mathematics course.

Table 3.7 Distribution of Points by Reporting Category—Algebra

	Form			
Reporting Category	**AR**	**BR**	**D**	**E**
Major Content	28	28	28	28
Additional & Supporting Content	13*	14	14	14
Expressing Mathematical Reasoning	11	11	11	11
Modeling & Application	15	15	15	15
Total	**67**	**68**	**68**	**68**

* A one-point item within the Additional & Supporting Content reporting category was dropped from scoring in Form AR.

Table 3.8 Distribution of Points by Reporting Category—Geometry

	Form			
Reporting Category	**AR**	**BR**	**D**	**E**
Major Content	26	26	26	26
Additional & Supporting Content	16	16	16	16
Expressing Mathematical Reasoning	11	11	11	11
Modeling & Application	15	15	15	15
Total	**68**	**68**	**68**	**68**

The Major Content category for mathematics is broken into subcategories by course as follows:

Table 3.9 Major Content Subcategories by Course

Course	Major Content Subcategories
Algebra	• Interpreting Functions • Solving Algebraically • Solving Graphically/Rate of Change
Geometry	• Congruence Transformations/Similarity • Similarity in Trigonometry/Modeling & Applying

The resulting 2020-2021 LEAP 2025 mathematics test blueprints are shown in Tables 3.10 and 3.11.

Table 3.10 Algebra I Test Blueprint

Reporting Category	Major Content	Additional & Supporting Content	Expressing Mathematical Reasoning	Modeling & Application
Task Type	Type I: I.1 (24 items, 24 points) I.2 (7 items, 14 points) I.4 (1 item, 4 points) Total: 32 items, 42 points (62% of total)		Type II: II.3 (1 item, 3 points) II.4 (2 items, 8 points) Total: 3 items, 11 points (16% of total)	Type III: III.3 (3 items, 9 points) III.6 (1 item, 6 points) Total 4 items, 15 points (22% of total)
Total OP Points	28 (41% of total)	14 (21% of total)	11 (16 % of total)	15 (22% of total)
Assessable Content	A1: A-APR.A.1 A1: A-CED.A.3 A1: A-CED.A.4 A1: A-REI.B.3 A1: A-REI.B.4a A1: A-REI.B.4b A1: A-REI.D.10 A1: A-REI.D.11 A1: A-REI.D.12 A1: A-SSE.A.1a A1: A-SSE.A.1b A1: A-SSE.A.2 A1: F-IF.A.1 A1: F-IF.A.2 A1: F-IF.B.4 A1: F-IF.B.5 A1: F-IF.B.6 LEAP.I.A1.1 LEAP.I.A1.2 LEAP.I.A1.3 LEAP.I.A1.4 LEAP.I.A1.5 LEAP.I.A1.6	A1: A-APR.B.3 A1: A-REI.C.6 A1: A-SSE.B.3a A1: A-SSE.B.3b A1: A-SSE.B.3c A1: F-BF.B.3 A1: F-IF.C.7a A1: F-IF.C.7b A1: F-IF.C.8a A1: F-IF.C.9 A1: F-LE.A.2 A1: S-ID.B.5 LEAP.I.A1.7	LEAP.II.A1.1 LEAP.II.A1.2 LEAP.II.A1.3 LEAP.II.A1.4 LEAP.II.A1.5 LEAP.II.A1.6 LEAP.II.A1.7 LEAP.II.A1.8 LEAP.II.A1.9 LEAP.II.A1.10	LEAP.III.A1.1 LEAP.III.A1.2 LEAP.III.A1.3 LEAP.III.A1.4

Table 3.11 Geometry Test Blueprint

Reporting Category	Major Content	Additional & Supporting Content	Expressing Mathematical Reasoning	Modeling & Application
Task Type	Type I: I.1 (24 items, 24 points) I.2 (7 items, 14 points) I.4 (1 item, 4 points) Total: 32 items, 42 points (62% of total)		Type II: II.3 (1 item, 3 points) II.4 (2 items, 8 points) Total: 3 items, 11 points (16% of total)	Type III: III.3 (3 items, 9 points) III.6 (1 item, 6 points) Total 4 items, 15 points (22% of total)
Total OP Points	26 (38% of total)	16 (24% of total)	11 (16% of total)	15 (22% of total)
Assessable Content	GM: G-CO.B.6 GM: G-GPE.B.6 GM: G-SRT.A.1a GM: G-SRT.A.1b GM: G-SRT.A.2 GM: G-SRT.B.5 GM: G-SRT.C.6 GM: G-SRT.C.7 GM: G-SRT.C.8 LEAP.I.GM.1 LEAP.I.GM.2	GM: G-C.A.2 GM: G-CO.A.1 GM: G-CO.A.3 GM: G-CO.A.5 GM: G-GMD.A.1 GM: G-GMD.A.3 GM: G-GMD.B.4 GM: G-GPE.A.1 LEAP.I.GM.3 LEAP.I.GM.4 LEAP.I.GM.5	LEAP.II.GM.1 LEAP.II.GM.2 LEAP.II.GM.3 LEAP.II.GM.4	LEAP.III.GM.1 LEAP.III.GM.2 LEAP.III.GM.3 LEAP.III.GM.4 LEAP.III.GM.5

Unlike the ELA test blueprints, which were organized by test sessions one through three, the mathematics test blueprints were organized by reporting categories, so it was necessary to define the general structure of the test forms by test session. The design goal was to have balanced test sessions with a variety of task types and equivalent testing times. For session 1a of the mathematics assessments, students were prohibited from using calculators, except those students with a calculator accommodation. Calculators were allowed to be used by all students in sessions 1b, 2, and 3. The general test structures (see Tables 3.12 and 3.13) guided test form sequencing and design. The LEAP 2025 Calculator Policy provided the basis for calculator designation of tasks and items.

Table 3.12 Algebra I Testing Sessions

Reporting Category	Session 1a: No Calculator	Session 1b: Calculator	Session 2: Calculator	Session 3: Calculator	Total
Major Content (points)	5	5	9	9	28
Additional & Supporting Content (points)	4	2	4	4	14
Expressing Mathematical Reasoning (points)	0	3	4	4	11
Modeling & Application (points)	0	3	6	6	15
Total Operational Points	9	13	23	23	68
Test Duration *(minutes)	25	55	80	80	240
# of Operational Items	I.1: 5 I.2: 2 I.4: 0 II.3: 0 II.4: 0 III.3: 0 III.6: 0	I.1: 3 I.2: 0 I.4: 1 II.3: 1 II.4: 0 III.3: 1 III.6: 0	I.1: 9 I.2: 2 I.4: 0 II.3: 0 II.4: 1 III.3: 0 III.6: 1	I.1: 7 I.2: 3 I.4: 0 II.3: 0 II.4: 1 III.3: 2 III.6: 0	I.1: 24 I.2: 7 I.4: 1 II.3: 1 II.4: 2 III.3: 3 III.6: 1

* Six Embedded Field Test Items were included throughout the assessment; they are included in the total time.

Table 3.13 Geometry Testing Sessions

Reporting Category	Session 1a: No Calculator	Session 1b: Calculator	Session 2: Calculator	Session 3: Calculator	Total
Major Content (points)	5	5	8	8	26
Additional & Supporting Content (points)	4	2	5	5	16
Expressing Mathematical Reasoning (points)	0	3	4	4	11
Modeling & Application (points)	0	3	6	6	15
Total Operational Points	9	13	23	23	68
Test Duration *(minutes)	25	55	80	80	240
# of Operational Items	I.1: 5 I.2: 2 I.4: 0 II.3: 0 II.4: 0 III.3: 0 III.6: 0	I.1: 3 I.2: 0 I.4: 1 II.3: 1 II.4: 0 III.3: 1 III.6: 0	I.1: 7 I.2: 3 I.4: 0 II.3: 0 II.4: 1 III.3: 0 III.6: 1	I.1: 9 I.2: 2 I.4: 0 II.3: 0 II.4: 1 III.3: 2 III.6: 0	I.1: 24 I.2: 7 I.4: 1 II.3: 1 II.4: 2 III.3: 3 III.6: 1

* Six Embedded Field Test Items were included throughout the assessment; they are included in the total time.

The following item types were used in the 2020-2021 LEAP 2025 mathematics assessments:

- Multiple-choice: This item type requires students to select one correct answer from four answer choices. It may appear as a one-part question, as part of a two-part question, or as part of a constructed-response item. The multiple-choice items are worth one point.
- Multiple select: This item type requires students to select more than one correct answer from more than four answer choices. It may appear as a one-part question, as part of a two-part question, or as part of a constructed-response item. The multiple select items are worth one point. Students must choose all correct answers and no incorrect answers to receive credit.
- Short answer: This item type requires students to enter a numeric response by typing from the keyboard. It may appear as a one-part question, as part of a two-part question, or as part of a constructed-response item. The short answer items are worth one point. Unless specified in the question, students will earn credit for an answer that is equivalent to the correct numerical answer. Proper rounding may be required. Answers to short answer items can be positive or negative and must be entered in integer or decimal form.
- Keypad input items: This item type requires students to enter a mathematical response using a customized pallet of numbers, operations, variables, and/or mathematical symbols; allows the use of all rational and irrational numbers, expressions, and equations; and scores all equivalent responses as correct unless noted otherwise. This item type may appear as a one-part question, as part of a two-part question, or as part of a constructed-response item.
- Constructed-response items: This item type requires students to respond to an open-ended question, which must be typed into a response box; students may use the equation builder tool (specific to the course) to insert mathematical characters. This item type can be a single- or multi-part item. Constructed-response items ask students to write explanations or justifications, model a process, and/or solve real-world, multistep contextual problems. Students may receive partial or full credit on constructed-response items, and maximum point values will vary by constructed-response task. Maximum values for constructed-response items are 3, 4, or 6 points.
- Technology enhanced items: This item type uses technology to capture student responses. Technology-enhanced items may appear as a one-part question, as part of a two-part question, or as part of a constructed-response item. The technology-enhanced items are worth one point. Technology-enhanced items may involve any of the following:
    - Bar graph: requires students to complete a bar graph or histogram by raising/lowering each bar to a value
    - Drag and drop: requires students to move draggable elements into one or more drop boxes
    - Drop-down menu: requires students to select from one or more drop-down menus to complete a sentence, phrase, or expression/equation/inequality
    - Hot spot: requires students to select one or more responses by choosing selectable areas on the screen
    - Match interaction table: requires students to select a checkbox in each row from two or more columns
    - Graph input: requires students to enter a response on a coordinate grid
    - Number line input: requires students to enter a response on a number line
    - Line plot: requires students to complete a line plot with “X” as the input

A variety of item types allows for the measurement of the full range of student performance.

The following table details the number of items by point value and task type and the number of points per task type for each form.

Table 3.14 Distribution of Mathematics Tasks and Points by Task Type

Form	Content Area	Type I				Type II			Type III			Total Number of Points
		1 pt. Tasks	2 pt. Tasks	4 pt. Tasks	Points	3 pt. Tasks	4 pt. Tasks	Points	3 pt. Tasks	6 pt. Tasks	Points	
AR	Algebra	23*	7	1	41	1	2	11	3	1	15	67*
BR	Algebra	24	7	1	42	1	2	11	3	1	15	68
D	Algebra	24	7	1	42	1	2	11	3	1	15	68
E	Algebra	24	7	1	42	1	2	11	3	1	15	68
AR	Geometry	24	7	1	42	1	2	11	1	2	15	68
BR	Geometry	24	7	1	42	1	2	11	1	2	15	68
D	Geometry	24	7	1	42	1	2	11	3	1	15	68
E	Geometry	24	7	1	42	1	2	11	3	1	15	68

* One Type I item worth 1 point was dropped from scoring in form AR.

Item Development and Selection

The processes of item development and selection are discussed in this section in compliance with the *Standards*.

Standard 4.7 states the following:

> The procedures used to develop, review, and try out items and to select items from the item pool should be documented (87).

The items used in the 2020-2021 LEAP 2025 high school ELA and mathematics assessments came from the New Meridian's item bank and the Louisiana-owned item bank.

The items selected for use on the 2020-2021 LEAP high school forms were used to equate to the LEAP 2025 scale, which is comparable to the PARCC scale. Operational forms were selected based on LEAP 2025 high school test blueprint specifications, which were supported by statistical data from New Meridian operational testing.

Considerations of Test Fairness in Item Development

Standard 3.2 is particularly relevant to fairness in item development:

> Test developers are responsible for developing tests that measure the intended construct and for minimizing the potential for tests being affected by construct-irrelevant characteristics, such as linguistic, communicative, cognitive, cultural, physical, or other characteristics (64).

Bias and sensitivity guidelines used to develop the New Meridian and Louisiana-owned items help ensure the assessments are fair for all groups of test takers, despite differences in characteristics that include, but are not limited to, disability status, ethnic group, race, gender, regional background, native language, religion, sexual orientation, and socioeconomic status. DRC relied strongly on the bias and sensitivity guidelines in the development of the assessments, particularly in item selection and review. To be included in the assessments, items had to comply with the bias and sensitivity guidelines and be approved by Louisiana educators involved in the Louisiana alignment and item review meetings.

New Meridian Item Reviews

As part of New Meridian's ongoing item development practices, several educator committees had already been convened to conduct rigorous reviews of every passage and item developed for the New Meridian assessment system prior to the items becoming a part of the item bank that included items and passages available for selection on Louisiana forms. These reviews include

- text reviews of all passages (during which participants review and edit passages independently and then discuss content and bias concerns as a grade-level group),
- item reviews (during which committees review and edit items for adherence to basic principles of universal design, accessibility guidelines, selected metadata fields, and a style guide),
- bias and sensitivity reviews (during which educators and community members review items and tasks to confirm the absence of issues relating to bias, fairness, and sensitivity to ensure that items and tasks do not unfairly advantage or disadvantage any student subgroup over another subgroup),
- editorial reviews (during which the review committee completes a copy edit review and records member comments), and
- data reviews (during which educators evaluate item-level statistics to determine eligibility of items and tasks to move forward to the operational assessments).

Additional information on New Meridian's item review processes and procedures can be found at the New Meridian Resource Center. Only items that have been approved by expert reviewers during text reviews (ELA only), item reviews, bias and sensitivity reviews, and editorial reviews are moved forward for field-testing. Of the field-tested items, only those determined to have acceptable statistics, either by having acceptable item parameters according to the data-review flagging criteria or by being approved by expert reviewers during data review, are eligible for review by Louisiana educators for potential use on an operational assessment. These processes follow the criteria set forth by the *Standards*.

Standard 3.1 states the following:

> Those responsible for test development, revision, and administration should design all steps of the testing process to promote valid score interpretations for intended score uses for the widest possible range of individuals and relevant subgroups in the intended population (63).

Standard 3.2 states the following:

> Test developers are responsible for developing tests that measure the intended construct and for minimizing the potential for tests being affected by construct-irrelevant characteristics, such as linguistic, communicative, cognitive, cultural, physical, or other characteristics (64).

Independent studies of New Meridian passages and items have found that the content being licensed assesses the skills that matter most and is rigorous, aligned to standards, and accessible to students with disabilities and English Learners. For more information on the studies performed, refer to New Meridian's website: https://resources.newmeridiancorp.org/research/.

3.2 Operational Test Selection

The operational tests administered in the 2020-2021 spring administration were the same forms used in the 2018-2019 spring administration. Therefore, information regarding operational test selection can be found in the 2018-2019 Technical Report.

3.3 Universal Design

Course-level assessments that follow universal design guidelines allow participation of the widest possible range of students, resulting in more valid inferences about students' performances. Such assessments may reduce the need for accommodations by reducing or eliminating access barriers associated with the tests themselves. Table 3.16 presents the elements of universal design (Thompson & Thurlow, 2002). The elements of universal design are relevant to both item development and form construction. This section describes how the elements of universal design were addressed in the construction of the test forms administered in 2020-2021 in compliance with AERA, APA, & NCME (2014) Standard 3.1, which states the following:

> Those responsible for test development, revision, and administration should design all steps of the testing process to promote valid score interpretations for intended score uses for the widest possible range of individuals and relevant subgroups in the intended population (63).

Universal design requires that assessments measure the performance of students with a wide range of abilities and skills, ensuring that students with diverse learning needs receive opportunities to demonstrate competence on the same content. To ensure that students can access the tests, the LEAP 2025 assessments include simple, clear, and intuitive instructions and procedures; maximum readability and comprehensibility; and maximum legibility. The online test specifications define how directions and test items are formatted online, including the spacing between an item stem and answer choices and other page elements (such as

online tools and Help files) to ensure consistent, clean visual appearance. Test directions at the beginning of each test session are clearly and simply stated, and the wording of such instructions is standardized as much as possible across tests to ensure clarity and consistency while being comparable to New Meridian.

Table 3.16 Elements of Universal Design

Element	Explanation
Inclusive Assessment Population	Tests designed for state, school system, or school accountability must include every student except those in the alternate assessment, and this is reflected in assessment design and field testing procedures.
Precisely Defined Constructs	The specific constructs tested must be clearly defined so that all construct-irrelevant cognitive, sensory, emotional, and physical barriers can be removed.
Accessible, Non-Biased Items	Accessibility is built into items from the beginning, and bias review procedures ensure that quality is retained in all items.
Amenable to Accommodations	The test design facilitates the use of needed accommodations (e.g., all items can be in braille form).
Simple, Clear, and Intuitive Instructions and Procedures	All instructions and procedures are simple, clear, and presented in understandable language.
Maximum Readability and Comprehensibility	A variety of readability and plain language guidelines are followed (e.g., sentence length and number of difficult words are kept to a minimum) to produce readable and comprehensible text.
Maximum Legibility	Characteristics that ensure easy decipherability are applied to text, tables, figures, illustrations, and response formats.

3.4 Accommodations and Designated Supports

AERA, APA, & NCME (2014) Standard 3.9 states the following:

> Test developers and/or test users are responsible for developing and providing test accommodations, when appropriate and feasible, to remove construct-irrelevant barriers that otherwise would interfere with examinees' ability to demonstrate their standing on the target constructs (67).

Students with disabilities, students with 504 plans, and English Learners (ELs) may be provided test administration accommodations based on their accommodation plan. More information on accommodations can be found in Chapter 4. Accommodation coding instructions can be found in the *Test Coordinator Manual*.

Accommodated print forms were developed for the high school ELA and mathematics tests for those students who were unable to participate in an online administration. For a detailed description of the process used to develop the accommodated print forms and how to modify technology-enhanced items for use in an accommodated print form, see Appendix A, "Accommodated Print Form Creation."

Braille forms were constructed for each course to enable students with visual impairments to participate in the LEAP 2025 assessments. Braille forms were based on the accommodated print forms. There are no large-

print versions of the accommodated print forms. Instead, students needing a large-print version use larger-sized monitors and/or the magnification features of the online testing system. All online test content has been developed to scale in relation to the available area on larger monitors while maintaining the correct aspect ratio. Specific recommendations on how to transcribe items into braille were provided by the braille publisher to produce the braille version of the LEAP 2025 high school assessments and the test administrator's notes that accompany the braille forms. The goal was to maximize the number of items on the braille forms that could be transcribed into braille.

The following assessment features were available to all students and do not require any documentation either prior to or during the assessment:

- blank scratch paper and graph paper
- calculators (to be used in the calculator section only)
- color overlay
- contrasting colors/reverse colors
- directions in native language
- equation builder
- bookmark
- general administration directions clarified
- general administration directions read aloud and repeated as necessary
- general masking
- headphones
- highlighters
- line guides
- magnifiers/variable zoom
- measurement tools
- redirection of student to the test
- specialized furniture or equipment
- sticky note/notepad
- strikethrough
- and writing/formatting tools (for ELA constructed-response items only).

Accessibility features were available for all students with the particular need documented in their Individualized Education Programs (IEPs), Individual Accommodation Plans (IAPs), English Learner (EL) plans, or Personal Needs Profiles (PNPs). The following accessibility features were available: individual testing, small group testing, student reads assessment aloud to himself or herself, adaptive and specialized equipment or furniture, and math read aloud (text-to-speech or human reader).

Accommodations were available for students who have an IEP, IAP, or EL plan. The following accommodations were available: braille test materials, calculation device and math tools for non-calculator sections of mathematics assessments, transferred answers, recorded answers, mathematics Spanish read aloud, translated mathematics test, and test read aloud (text-to-speech). For details on these accessibility features and accommodations, see the *LEAP 2025 Accommodations and Accessibility Features User Guide*.

For a detailed description of the process used to develop the Spanish translation forms of the mathematics tests, see Appendix B, "Forms Development Process for Spanish Translations Forms."

3.5 Item and Task Specifications

AERA, APA, & NCME (2014) Standard 4.12 states the following:

> Test developers should document the extent to which the content domain of a test represents the domain defined in the test specifications (89).

The item and task specifications are designed to ensure that the assessment items measure the assessment's claims. The purpose of the item and task specifications is to define the characteristics of the items and tasks that will provide the evidence to support one or more claims. To do this, the item and task specifications delineate the types of evidence, or targets, that should be elicited for each reporting category within a grade level. The specifications provide explicit guidance on how to write items to elicit the desired evidence.

The item and task specifications provide guidance on how to measure the targets (i.e., standards) first found in the content specifications and guidelines on how to create the items that are specific to each assessment target and reporting category. In ELA and mathematics, item specifications describe the knowledge, skills, and processes being measured by each item type aligned to particular standards.

These item specifications were developed for each course and standard to delineate the expectations of knowledge and skill to be included on test questions. In addition, the ELA and mathematics item and stimulus specifications provide guidance on determining the appropriateness of task and stimulus materials (i.e., the materials that a student must refer to when working on a test question). The stimulus specifications also provide information on the characteristics of stimuli or activities that should be avoided because they are not important to the knowledge, skill, or process being measured. This underscores DRC's efforts to select items that are accessible to the widest range of students possible; in other words, 2020-2021 LEAP 2025 items were selected according to the elements of universal design.

3.6 Summary

In summary, the overall purpose of this chapter is to explicate the procedures used in the development of the forms administered during the 2020-2021 LEAP 2025 high school administrations. The efforts by LDOE and DRC in developing the LEAP 2025 high school assessments are in alignment with multiple best practices of the test industry but, in particular, support the following AERA, APA, & NCME (2014) standards:

> *Standard 3.1* Those responsible for test development, revision, and administration should design all steps of the testing process to promote valid score interpretations for intended score uses for the widest possible range of individuals and relevant subgroups in the intended population (63).
>
> *Standard 3.2* Test developers are responsible for developing tests that measure the intended construct and for minimizing the potential for tests being affected by construct-irrelevant characteristics, such as linguistic, communicative, cognitive, cultural, physical, or other characteristics (64).
>
> *Standard 3.9* Test developers and/or test users are responsible for developing and providing test accommodations, when appropriate and feasible, to remove construct-irrelevant barriers that otherwise would interfere with examinees' ability to demonstrate their standing on the target constructs (67).

Standard 4.0 Tests and testing programs should be designed and developed in a way that supports the validity of interpretations of the test scores for their intended uses. Test developers and publishers should document steps taken during the design and development process to provide evidence of fairness, reliability, and validity for intended uses for individuals in the intended examinee population (85).

Standard 4.1 Test specifications should describe the purpose(s) of the test, the definition of the construct or domain measured, the intended examinee population, and interpretations for intended uses. The specifications should include a rationale supporting the interpretations and uses of test results for the intended purpose(s) (85).

Standard 4.7 The procedures used to develop, review, and try out items and to select items from the item pool should be documented (87).

Standard 4.12 Test developers should document the extent to which the content domain of a test represents the domain defined in the test specifications (89).

Chapter 4: Test Administration

Chapter 4 of the technical report describes the processes implemented and the information disseminated to help ensure standardized test administration procedures and, thus, uniform test administration conditions for students. According to the *Standards for Educational and Psychological Testing* (American Educational Research Association [AERA], American Psychological Association [APA], & National Council on Measurement in Education [NCME], 2014), "The usefulness and interpretability of test scores require that a test be administered and scored according to the test developer's instructions" (111). This chapter examines how test administration procedures implemented for the 2020–2021 Louisiana Education Assessment Program (LEAP 2025) strengthen and support the intended score interpretations and reduce construct-irrelevant variance that could threaten the validity of score interpretations.

Chapter 4 demonstrates how the LEAP 2025 assessments adhere to AERA, APA, & NCME (2014) Standards 4.15, 6.1, 6.2, 6.3, 6.4, 6.6, and 6.7. Each standard will be explicated in the relevant section of this chapter.

To ensure that the LEAP 2025 assessments are administered in accordance with the department's mandates, the LDOE takes a primary role in communicating with and training school system personnel. The development of the assessments is a collaborative effort between LDOE and DRC. The LDOE conveys to school systems the purpose of the assessments and the importance of test administration being consistent with test industry standards. The tests and administration standards must also meet the State Board of Elementary and Secondary Education policies and the mandates of both state and federal legislation.

To accomplish these goals, the LDOE provides train-the-trainer opportunities for school system test coordinators, who, in turn, administer test-administration training to schools within their school systems. The LDOE conducts quality assurance visits during testing to ensure that school systems adhere to the standardized administration of the tests.

The school system test coordinators are responsible for the schools within their school systems. They disseminate information to each school, assist with test administration, and serve as liaisons between the LDOE and the schools in their system. The LDOE also provides assistance with and interpretation of assessment data and test results.

Ancillary materials for the LEAP 2025 test administration contribute to the body of evidence of the validity of score interpretation. This section examines how the test materials address the standards related to test administration procedures.

For the administration of the LEAP 2025 High School assessments, DRC produced the following test administration manuals (TAMs): *High School Test Administration Manual: LEAP 2025, Fall 2020; High School Test Administration Manual, Spring 2021; High School Test Administration Manual, Summer 2021.* DRC also produced the following test coordinator manuals (TCMs): *Test Coordinator Manual: LEAP 2025, Fall 2020; Test Coordinator Manual: LEAP 2025, Spring 2021; Test Coordinator Manual: LEAP 2025, Summer 2021*. LDOE assessment administration and development staff review these manuals, provide feedback, and give final approval. Each TCM includes information about LEAP 2025 HS ELA, mathematics, U.S. history, and biology. It provides detailed instructions for school system and school test coordinators on distributing and collecting test materials and for returning them to DRC as outlined in its table of contents.

Test Coordinator Manual Table of Contents

The TAMs provide detailed instructions for administering the LEAP 2025 assessments. The manuals include instructions for test security, test preparation, administration of tests, and post-test procedures. Information included in the TAMs is listed below.

Test Administrators Manual Table of Contents

15. Index

The *Standards* contain multiple references that are relevant to test administration. Information in the TAMs addresses these standards.

The directions for test administration found in the manual address Standard 4.15, which states:

> The directions for test administration should be presented with sufficient clarity so that it is possible for others to replicate the administration conditions under which the data on reliability, validity, and (where appropriate) norms were obtained. Allowable variations in administration procedures should be clearly described. The process for reviewing requests for additional testing variations should also be documented (90).

The LEAP 2025 Test Administration Manuals provide instructions for activities conducted before, during, and after testing with sufficient detail and clarity to support reliable test administrations by qualified test administrators. To ensure uniform administration conditions throughout the state, instructions in the manuals describe the following: general rules of online testing; assessment duration, timing, and sequencing information; and the materials required for testing.

Furthermore, the standardized procedures addressed in the test administration manual need to be followed. The *Standards* state in Standard 6.1:

> "Test administrators should follow carefully the standardized procedures for administration and scoring specified by the test developer and any instructions from the test user" (114).

It was essential that the LEAP 2025 was administered according to the prescribed test administration manual to ensure the usefulness and interpretability of the test scores and to minimize sources of construct-irrelevant variance. It should be noted that adhering to the test schedule is also a critical component. The test administration manuals include instructions for scheduling the test within the state testing window. The test administration manuals also contain the schedule for timing each test session. The test timing schedule is presented in Table 4.1.

> *Standard 6.3* Changes or disruptions to standardized test administration procedures or scoring should be documented and reported to the test user (115).

The LDOE staff administer reports on testing concerns that describe a wide range of improper activities that may occur during testing, including the following: copying and reviewing test questions with students; cueing students during testing, verbally or with written materials on the classroom walls; cueing students nonverbally, such as by tapping or nodding the head; allowing students to use a calculator on parts of the test where it is not allowed; allowing students to correct or complete answers after tests have been submitted; splitting sessions into two parts; ignoring the standardized directions in the online assessment; reading the ELA assessment to students (with the exception of those students with the read-aloud accommodation); paraphrasing parts of the test to students; changing or completing (or allowing other school personnel to change or complete) student answers; allowing accommodations that are not written in the Individualized Education Program (IEP); allowing accommodations for students who do not have an IEP; or defining terms on the test.

Each administration includes an administrative error retest, which provides an opportunity for students to retake a test that was voided during the regular test window because of improper activities that occurred during testing (e.g., the student was not given enough time to complete the test, the student was not

provided proper accommodations during the testing time, the teacher or administrator provided information or answers that resulted in the test being voided).

> *Standard 6.4* The testing environment should furnish reasonable comfort with minimal distractions to avoid construct-irrelevant variance (116).

The test administration manuals outline the steps that teachers should take to prepare classroom environments for administering the LEAP 2025 assessments. These steps include the following:

- Determine the layout of the classroom environment.
- Plan seating arrangements. Allow enough space between students to prevent the sharing of answers.
- Eliminate distractions such as bells or telephones.
- Use a Do Not Disturb sign on the door of the testing room.
- Make sure classroom maps, charts, and any other materials that relate to the content and processes of the test are covered, removed, or out of students' view.

> *Standard 6.6* Reasonable efforts should be made to ensure the integrity of test scores by eliminating opportunities for test takers to attain scores by fraudulent or deceptive means (116).

The test administration manuals present instructions for post-test activities to ensure that online tests are submitted and that printed test materials are handled properly to maintain the integrity of student information and test scores. Detailed instructions guide test examiners in submitting all online test records. For students who were administered a braille test form, examiners are instructed to transcribe students' responses from the braille test form into the online testing system (INSIGHT) exactly as the responses appear in the original form.

> *Standard 6.7* Test users have the responsibility of protecting the security of test materials at all times (117).

Throughout the manuals, test coordinators and examiners are reminded of test security requirements and procedures to maintain test security. Specific actions that are direct violations of test security are so noted. Detailed information about test security procedures is presented under "Test Security" in the test administration manuals.

4.1 Return Material Forms and Guidelines

The test coordinator manual instructs test coordinators on how to organize, pack, and return testing materials to DRC for secure inventory purposes. The LDOE assessment administration and development staff have opportunities to review these materials, provide feedback, and give final approval. The purpose of the instructions is to ensure that the secure test materials are properly accounted for and organized appropriately for return shipment.

4.2 Security Checklists

As soon as printed test materials are received by a school system, the district test coordinator ensures the first and last security barcodes on the tests match the packing list they received. The district test coordinator then packages the test materials to be sent to schools. District test coordinators are required to return communication assistance scripts (CAS) and braille test materials to DRC. School systems are required to document nonstandard situations, including lost, damaged, destroyed, extra, or missing materials. Any

material not accounted for is placed on a missing materials list, which is used by DRC and LDOE to follow up with all districts to ensure security of all materials.

4.3 Interpretive Guides

An understanding of what test scores mean and how to interpret score reports is essential to making valid interpretations of the test scores. The *LEAP 2025 HS Interpretive Guide* is written for Louisiana teachers and administrators who receive the LEAP 2025 score reports. More details about the guide can be found in Chapter 7.

4.4 Test Security Measures

Maintaining the security of all test materials is crucial to preventing the possibility of random or systematic errors, such as unauthorized exposure of test items, that would affect the valid interpretation of test scores. Several test security measures are implemented for the LEAP 2025 assessments. Test security procedures are discussed throughout the Test Coordinator Manuals and Test Administration Manuals.

Test coordinators and administrators are instructed to keep all test materials in locked storage, except during actual test administration, and access to secure materials must be restricted to authorized individuals only (e.g., test administrators and the school test coordinator). During testing sessions, the test administrators are directly responsible for the security of the LEAP 2025 assessments, must account for all test materials, and supervise the test administration at all times.

Data Forensic Analyses

Due to the importance of the LEAP 2025 assessments, it is prudent to ensure that the results from the assessments are based on effective instruction and true student achievement. While there are many ways to achieve meaningful understanding of student knowledge via test scores, there are also ways to obtain higher test scores that are not related to actual learning. To assist in ensuring that assessment results are valid, data forensic analyses are conducted to help separate meaningful gains from spurious gains. It is important to note that although the results may be used to identify potential problems within a school, the identification of a problem is not an accusation of misconduct.

Multiple methods of analysis were incorporated into the forensic analysis. The following methods were applied:

- Response-Change Analysis
- Score-Fluctuation Analysis
- Item Exposure Monitoring
- Web Monitoring
- Plagiarism Detection

Response-Change Analysis

Students make changes to answer choices when taking the LEAP 2025, and this is expected behavior. Unfortunately, changing student answers is also an opportunity for school personnel to improve classroom performance. The response-change analysis focuses on identifying school- and test-administrator level response-change patterns that are statistically improbable when compared to the expected pattern at the state level.

Score-Fluctuation Analysis

It is anticipated that performance on the LEAP 2025 will improve over time from legitimate sources such as changes in the curriculum and improvement in instruction. However, large and unexpected score changes

may be a sign of testing impropriety. The LDOE applied an approach wherein the state's change in performance from one year to the next is compared to a schools' and test administrators' change in performance during the same time frame. Schools and test administrators were identified when the level of change was statistically unexpected.

Item Exposure Monitoring

The Fall 2020 test administration included two testing windows; there was a testing window in December 2020 and a testing window in January 2021. Due to the same form being used in both windows, item performance was examined in the second window to ensure that item content had not been exposed. In addition to reviewing fit plots for good alignment of an item's performance across the windows, an item's moving *p*-value and point-biserial averages were produced daily. During the January testing window, If an item's moving average *p*-value was larger than expected compared to the previous day's or the December average, the item was flagged. Similar methodology was also applied in the spring 2021 test window due to the reuse of the spring 2019 test forms.

Web Monitoring

LEAP 2025 operational test content should not appear outside the boundaries of the forms administered. To protect Louisiana test content, the internet is monitored for postings which contain, or appear to contain, potentially exposed and/or copied LDOE test content. When test content is verified, steps are taken so that the infringing content is removed quickly.

Plagiarism Detection

The LDOE monitors for two different plagiarism situations: copying from student to student and copying from an outside source, such as Wikipedia or other internet sources. Instances of plagiarism are identified regardless of whether an item is scored by human scorers or artificial intelligence. Alerts are set to identify responses that may indicate teacher interference, plagiarism, or disturbing content (e.g., possible physical or emotional abuse, suicidal ideation, threats of harm to the student in question or others, etc.). Alerted responses are given additional review so the appropriate response can be taken.

4.5 Test Administration

The 2020–2021 assessments were administered to students within the state testing windows of December 1 through 18, 2020, or January 6-26, 2021; April 15 through May 21, 2021; and June 21–25, 2021. Each session of the LEAP 2025 assessments was required to be administered in one block of time.

Time

All sessions of the LEAP 2025 high school ELA and mathematics assessments were timed. Only students with an extended time accommodation were permitted to exceed the established time limits of any given session. The timing schedule of the LEAP 2025 assessments is presented in Table 4.1.

Table 4.1 LEAP 2025 Administration Schedule Timing by Session

Course	Session	Minutes
English I	1	90
	2	90
	3	80
English II	1	90
	2	90
	3	80
Algebra I	1a	25
	1b	55
	2	80
	3	80
Geometry	1a	25
	1b	55
	2	80
	3	80

Accommodations

Accommodations are allowed on the LEAP 2025 assessments.

Accommodations may be used by a student who qualifies under the Individual with Disabilities Act (IDEA), has an IEP or a Section 504 plan of the Americans with Disabilities Act, or identifies as an English Learner (EL). Accommodations must be specified in the qualifying student's individual plan and must be consistent with accommodations used during daily classroom instruction and testing. The use of any accommodation must be indicated on the student information sheet at the time of test administration. AERA, APA, & NCME Standard 6.2 states:

> When formal procedures have been established for requesting and receiving accommodations, test takers should be informed of these procedures in advance of testing (115).

In compliance with this standard, the LEAP 2025 Test Administration Manuals contain the list of universal tools, designated supports, and accommodations permissible for the LEAP 2025 assessments. Further guidance can be found in the LEAP 2025 Accommodations and Accessibility Features User Guide.

Visually impaired students may be provided braille forms for any assessment.

Tables 4.2 through 4.4 summarize the numbers of reportable students receiving accommodations or designated features by type for the 2020-2021 LEAP 2025 HS administrations. Accommodation assignment guidance is provided in the *LEAP 2025 Accommodations and Accessibility Features User Guide*. The analyses

are based on census data and include only those students who received accommodations or designated features and received a scale score on the ELA or mathematics LEAP 2025 high school assessments. The percentage represents the percentage of the census population receiving that accommodation or designated feature.

Table 4.2 Fall 2020 Number and Percentage of Students Receiving Accommodations by Accommodation/Designated Feature Type, as identified in DRC INSIGHT Portal (eDIRECT)

Content	Accommodation/Designated Feature	Number	Percentage
English I	Text-to-Speech	≥830	11.80
	Accommodated Print	<50	NR
	Human Read Aloud	≥110	1.57
	Native Language Word-to-Word Dictionary	≥380	5.37
	Directions in Native Language	≥100	1.53
	Communication Assistance	<50	NR
	Transferred Answers	<50	NR
	Answers Recorded	<50	NR
	Extended Time	≥1,870	26.45
	Individual/Small Group Administration	≥1,110	15.80
	Braille	<50	NR
English II	Text-to-Speech	≥640	7.22
	Accommodated Print	<50	NR
	Human Read Aloud	≥120	1.35
	Native Language Word-to-Word Dictionary	≥450	5.12
	Directions in Native Language	≥110	1.32
	Communication Assistance	<50	NR
	Transferred Answers	<50	NR
	Answers Recorded	<50	NR
	Extended Time	≥1,910	21.54
	Individual/Small Group Administration	≥990	11.18
	Braille	<50	NR
Algebra I	Text-to-Speech	≥590	11.67
	Accommodated Print	<50	NR
	Human Read Aloud	≥80	1.70
	Native Language Word-to-Word Dictionary	≥190	3.74
	Directions in Native Language	<50	NR
	Communication Assistance	<50	NR
	Transferred Answers	<50	NR
	Answers Recorded	<50	NR
	Calculator	≥510	10.16
	Extended Time	≥1,150	22.70
	Individual/Small Group Administration	≥680	13.49
	Braille	<50	NR

Table 4.3 Fall 2020 Number and Percentage of Students Receiving Accommodations by Accommodation/Designated Feature Type, as identified in DRC INSIGHT Portal (eDIRECT) (Continued)

Content	Accommodation/Designated Feature	Number	Percentage
Geometry	Text-to-Speech	≥370	6.63
	Accommodated Print	<50	NR
	Human Read Aloud	<50	NR
	Native Language Word-to-Word Dictionary	≥180	3.29
	Directions in Native Language	≥50	0.91
	Communication Assistance	<50	NR
	Transferred Answers	<50	NR
	Answers Recorded	<50	NR
	Calculator	≥270	4.77
	Extended Time	≥830	14.70
	Individual/Small Group Administration	≥440	7.82
	Braille	<50	NR

Table 4.4 Spring 2021 Number and Percentage of Students Receiving Accommodations by Accommodation/Designated Feature Type, as identified in DRC INSIGHT Portal (eDIRECT)

Accommodation/Designated Feature Type: Spring 2021			
Content	**Accommodation/Designated Feature**	**Number**	**Percentage**
English I	Text-to-Speech	≥4,390	9.39
	Accommodated Print	<50	NR
	Human Read Aloud	≥450	0.98
	Native Language Word-to-Word Dictionary	≥1,200	2.59
	Directions in Native Language	≥250	0.55
	Communication Assistance	<50	NR
	Transferred Answers	≥50	0.11
	Answers Recorded	≥70	0.16
	Extended Time	≥10,150	21.72
	Individual/Small Group Administration	≥5,680	12.17
	Braille	<50	NR
English II	Text-to-Speech	≥3,500	8.59
	Accommodated Print	<50	NR
	Human Read Aloud	≥380	0.94
	Native Language Word-to-Word Dictionary	≥810	2.00
	Directions in Native Language	≥170	0.44
	Communication Assistance	<50	NR
	Transferred Answers	<50	NR
	Answers Recorded	≥50	0.14
	Extended Time	≥8,100	19.84
	Individual/Small Group Administration	≥4,510	11.06
	Braille	<50	NR
Algebra I	Text-to-Speech	≥5,730	11.82
	Accommodated Print	<50	NR
	Human Read Aloud	≥560	1.16
	Native Language Word-to-Word Dictionary	≥1,030	2.13
	Directions in Native Language	≥200	0.43
	Communication Assistance	<50	NR
	Transferred Answers	≥50	0.11
	Answers Recorded	≥70	0.15
	Calculator	≥5,210	10.76
	Extended Time	≥10,330	21.31
	Individual/Small Group Administration	≥5,970	12.32
	Braille	<50	NR

Accommodation/Designated Feature Type: Spring 2021 (continued)			
Content	**Accommodation/Designated Feature**	**Number**	**Percentage**
Geometry	Text-to-Speech	≥2,200	6.43
	Accommodated Print	<50	NR
	Human Read Aloud	≥180	0.55
	Native Language Word-to-Word Dictionary	≥430	1.26
	Directions in Native Language	≥100	0.31
	Communication Assistance	<50	NR
	Transferred Answers	<50	NR
	Answers Recorded	<50	NR
	Calculator	≥2,010	5.89
	Extended Time	≥4,790	14.00
	Individual/Small Group Administration	≥2,620	7.65
	Braille	<50	NR

Table 4.5 Summer 2021 Number and Percentage of Students Receiving Accommodations by Accommodation/Designated Feature Type, as identified in DRC INSIGHT Portal (eDIRECT)

Accommodation/Designated Feature Type: Summer 2021			
Content	**Accommodation/Designated Feature**	**Number**	**Percentage**
English I	Text-to-Speech	≥690	27.01
	Accommodated Print	<50	NR
	Human Read Aloud	≥150	5.93
	Native Language Word-to-Word Dictionary	≥170	6.94
	Directions in Native Language	≥90	3.68
	Communication Assistance	<50	NR
	Transferred Answers	<50	NR
	Answers Recorded	<50	NR
	Extended Time	≥1,180	46.07
	Individual/Small Group Administration	≥610	23.98
	Braille	<50	NR
English II	Text-to-Speech	≥650	24.64
	Accommodated Print	<50	NR
	Human Read Aloud	≥110	4.24
	Native Language Word-to-Word Dictionary	≥110	4.43
	Directions in Native Language	<50	NR
	Communication Assistance	<50	NR
	Transferred Answers	<50	NR
	Answers Recorded	<50	NR

Accommodation/Designated Feature Type: Summer 2021			
Content	**Accommodation/Designated Feature**	**Number**	**Percentage**
	Extended Time	≥1,050	39.89
	Individual/Small Group Administration	≥510	19.61
	Braille	<50	NR
Algebra I	Text-to-Speech	≥870	24.80
	Accommodated Print	<50	NR
	Human Read Aloud	≥170	4.80
	Native Language Word-to-Word Dictionary	≥100	3.08
	Directions in Native Language	≥50	1.64
	Communication Assistance	<50	NR
	Transferred Answers	<50	NR
	Answers Recorded	<50	NR
	Calculator	≥810	23.07
	Extended Time	≥1,250	35.53
	Individual/Small Group Administration	≥630	17.79
	Braille	<50	NR
Geometry	Text-to-Speech	≥160	15.78
	Accommodated Print	<50	NR
	Human Read Aloud	<50	NR
	Native Language Word-to-Word Dictionary	<50	NR
	Directions in Native Language	<50	NR
	Communication Assistance	<50	NR
	Transferred Answers	<50	NR
	Answers Recorded	<50	NR
	Calculator	≥150	14.74
	Extended Time	≥250	23.82
	Individual/Small Group Administration	≥130	12.67
	Braille	<50	NR

4.6 Summary

In summary, the overall purpose of each of the test administration trainings and the ancillary materials is to keep school systems informed about policies and procedures related to testing in general and the LEAP 2025 program in particular. The information imparted is clearly related to standardizing the administration of the LEAP 2025, maintaining the security of the assessment, allowing access to the assessments for special populations through appropriate accommodations, and maintaining the integrity of the scores. These communication and training efforts by LDOE and the ancillary information developed by DRC address multiple best practices of the testing industry but, in particular, are related to the following standards:

> *Standard 4.15* The directions for test administration should be presented with sufficient clarity so that it is possible for others to replicate the administration conditions under which the data on reliability, validity, and (where appropriate) norms were obtained. Allowable variations in administration procedures should be clearly described. The process for reviewing requests for additional testing variations should also be documented (90).

Standard 6.1 Test administrators should follow carefully the standardized procedures for administration and scoring specified by the test developer and any instructions from the test user (114).

Standard 6.3 Changes or disruptions to standardized test administration procedures or scoring should be documented and reported to the test user (115).

Standard 6.4 The testing environment should furnish reasonable comfort with minimal distractions to avoid construct-irrelevant variance (116).

Standard 6.6 Reasonable efforts should be made to ensure the integrity of test scores by eliminating opportunities for test takers to attain scores by fraudulent or deceptive means (116).

Standard 6.7 Test users have the responsibility of protecting the security of test materials at all times (117).

Chapter 5: Scoring of Constructed-Response and Technology-Enhanced Items

In this chapter, the scoring process used for the 2020-2021 LEAP 2025 high school ELA and mathematics assessments is described, with a particular focus on the handscoring of constructed-response items and the automated scoring of technology-enhanced items. At the end of this section, the results of the inter-rater reliability for the handscoring of the 2020-2021 LEAP 2025 constructed-response items are presented.

Chapter 5 demonstrates how the LEAP 2025 assessments adhere to the American Educational Research Association, American Psychological Association, & National Council on Measurement in Education (AERA, APA, & NCME, 2014) Standards 4.18, 4.20, 6.8, and 6.9. Each standard is presented in the pertinent section of this chapter. Standard 4.18 provides some general guidance for Chapter 5:

> Procedures for scoring and, if relevant, scoring criteria, should be presented by the test developer with sufficient detail and clarity to maximize the accuracy of scoring. Instructions for using rating scales or for deriving scores obtained by coding, scaling, or classifying constructed responses should be clear. This is especially critical for extended-response items such as performance tasks, portfolios, and essays (91).

Chapter 5 explains the procedures used for scoring the LEAP 2025 ELA and mathematics constructed-response items and technology-enhanced items. The scoring criteria used for each item are not presented in this chapter to preserve the integrity of the items for future use.

5.1 Constructed-Response Item Scoring Process

Constructed-response items were scored by human raters who were trained by DRC. Handscoring and Artificial Intelligence (AI) processing rules are detailed in Appendix C. Some ELA items across English I and English II (noted in Table 5.1) were scored by an AI engine, Pearson's Intelligent Essay Assessor (IEA), using scoring models previously developed by Pearson. Second reads of 10% of these responses were completed by human scorers; handscoring supervisors also reviewed the responses that IEA was not able to score.

Table 5.1 Constructed-Response Operational Scoring

Administration	Course	Handscoring Only	AI Scoring	AI Vendor
Fall 2020	English I	N/A	906152, 902161	Pearson
	English II	902354	902331	Pearson
	Algebra I	All CRs	N/A	
	Geometry	All CRs	N/A	
Spring 2021	English I	N/A	983215, 914552	Pearson
	English II	N/A	983688, 983642	Pearson
	Algebra I	All CRs	N/A	
	Geometry	All CRs	N/A	
Summer 2021	English I	N/A	902161, 902152	Pearson
	English II	N/A	902331, 906197	Pearson
	Algebra I	All CRs	N/A	
	Geometry	All CRs	N/A	

Selection of Scoring Evaluators

Standard 4.20 states the following:

> The process for selecting, training, qualifying, and monitoring scorers should be specified by the test developer. The training materials, such as the scoring rubrics and examples of test takers' responses that illustrate the levels on the rubric score scale, and the procedures for training scorers should result in a degree of accuracy and agreement among scorers that allows the scores to be interpreted as originally intended by the test developer. Specifications should also describe processes for assessing scorer consistency and potential drift over time in raters' scoring (92).

The following sections explain how scorers were selected and trained for the LEAP 2025 ELA and mathematics handscoring process and how scorers were monitored throughout the handscoring process.

The Recruitment and Interview Process

DRC strives to develop a highly qualified, experienced core of evaluators to appropriately maintain the integrity of all projects. All readers hired by DRC to score the 2020-2021 LEAP 2025 high school ELA and mathematics test responses had at least a four-year college degree.

DRC has a human resources director dedicated solely to recruiting and retaining the handscoring staff. Applications for reader positions are screened by the handscoring project manager, the human resources director, or recruiting staff to create a large pool of potential readers. In the screening process, preference is given to candidates with previous experience scoring large-scale assessments and with degrees emphasizing the appropriate content areas. At the personal interview, reader candidates are asked to demonstrate their proficiency in writing by responding to a DRC writing topic and their proficiency in mathematics by solving word problems with correct work shown. These steps result in a highly qualified and diverse workforce. DRC personnel files for readers and team leaders include evaluations for each project completed. DRC uses these evaluations to place individuals on projects that best fit their professional backgrounds, their college degrees, and their performances on similar projects at DRC. Once placed, all readers go through rigorous training and qualifying procedures specific to the project on which they are placed. Any scorer who does not complete this training and demonstrate the ability to apply the scoring criteria by qualifying at the end of the process is not allowed to score live student responses.

Security

Whether training and scoring are conducted within a DRC facility or done remotely, security is essential to our handscoring process. When users log into DRC's secure, web-based scoring application, ScoreBoard, they are required to read and accept our security policy before they are allowed to access any project. For each project, scorers are also required to read and sign non-disclosure agreements, and during training emphasis is always given to what security means, the importance of maintaining security, and how this is accomplished.

Readers only have access to student responses they are qualified to score. Each scorer is assigned a unique username and password to access DRC's imaging system and must qualify before viewing any live student responses. DRC maintains full control of who may access the system and which item each scorer may score. No demographic data is available to scorers at any time.

Handscoring Training Process

Standard 6.9 specifies:

> Those responsible for test scoring should establish and document quality control processes and criteria. Adequate training should be provided. The quality of scoring should be monitored and documented. Any systematic source of scoring errors should be documented and corrected (118).

Training Material Development

DRC scoring supervisors trained scorers using training materials from two sources.

1. PARCC-approved training materials provided by New Meridian. These materials were developed according to processes described in PARCC technical reports and include the following:

    - Passages, prompts, and associated stimuli
    - Rubrics
    - Anchor sets
    - Practice sets
    - Qualifying sets (for prototype items only)

2. Math training materials developed by DRC in conjunction with and approved by LDOE. These materials were made for use with DRC-developed math items (which were newly operational in the spring of 2019) according to processes described in DRC's response to the LDOE's "REQUEST FOR PROPOSALS For LEAP 2025 Assessment Administration (RFP #: 815200-20150723001)".

    - Prompts
    - Rubrics
    - Anchor sets
    - Practice sets
    - Qualifying sets (for all DRC-developed items)

Training and Qualifying Procedures

Handscoring involves training and qualifying team leaders and evaluators, monitoring scoring accuracy and production, and ensuring security of both the test materials and the scoring facilities. LDOE reviews training

materials and oversees the training process. An explanation of the training and qualification procedures follows.

DRC used the PARCC-approved mathematics and ELA training and qualifying materials to score two categories of items: "prototype" items and "abbreviated" items. Note that, like the PARCC "prototype" items for math, full sets of training and qualifying materials were also developed for all DRC-developed math items. The training and qualifying procedures DRC used for these items was the same process outlined below for PARCC-approved "prototype" math items.

Prototype Items

A small number of items (two each for Algebra I and Geometry and one for ELA) included in the Louisiana forms were prototype items, meaning they had full sets of associated training materials, including anchor sets, practice sets, and qualifying sets. DRC started the training process with a review of passages and items, rubrics, and anchor sets, followed by the scoring and discussion of practice sets and qualifying sets. Once this process was completed for a prototype item included on the Louisiana form, qualified readers started scoring live student responses for that item.

Abbreviated Items

Abbreviated items required a two-step training and qualifying process. First, scorers trained and qualified as described above using PARCC-approved materials for an associated prototype item that was similar to the abbreviated one they would be scoring on the Louisiana form.[1] Readers who did not qualify on the prototype item training were not allowed to continue the training.

After qualifying on the associated prototype item training, readers received additional item-specific training on the abbreviated item they were going to score. This consisted of an item-specific anchor set and two item-specific practice sets. After completing the abbreviated item training, readers could begin scoring live student responses for the abbreviated item.

[1] *Item associations were determined by PARCC/Pearson with the understanding that aspects of training are generalizable across similar items. For mathematics, the determination of prototype versus abbreviated items was made by PARCC and Pearson based on similar item types and evidence statements. For ELA items, this determination by PARCC and Pearson was based on grade and task type.*

The following tables detail the composition of the training materials provided by New Meridian for mathematics and ELA.

Table 5.2 Mathematics Training Set Composition

Set Type	Prototype Item Training	Abbreviated Item Training	Annotated
Anchor Set	3 responses per score point (Composite items had 3 responses per composite score.)	3 responses per score point (Composite items had 3 responses per composite score.)	Yes
Practice Set 1	10 responses representing the range of responses	10 responses representing the range of responses	Yes
Practice Set 2	10 responses representing the range of responses	10 responses representing the range of responses	Yes
Qualifying Set 1	10 responses comparable to the anchor set responses		No
Qualifying Set 2	10 responses comparable to the anchor set responses		No
Qualifying Set 3	10 responses comparable to the anchor set responses		No
For DRC-developed math items, examples of responses at the top score points may not have been present in some anchor, training, and qualifying sets as there were few or no examples found during rangefinding or subsequent field test scoring. In such cases, DRC Scoring Directors identified examples of these scores during live scoring to supplement reader training.			

Table 5.3 ELA Training Set Composition

Set Type	Prototype Item Training	Abbreviated Item Training	Annotated
Anchor Set*	3 responses per score point	16 responses per item: • Anchor Sets for abbreviated RST and LAT item training included scores for the combined trait Reading Comprehension and Written Expression (RCWE). • Anchor Sets for abbreviated NWT item training included scores for Written Expression (WE).	Yes
Practice Set 1	5 responses representing the range of responses for ● the Reading Comprehension and Written Expression (RCWE) trait (for LAT and RST items) ● the Written Expression trait (for NWT items)	10 responses representing the range of responses for the trait appropriate to the task type	Yes
Practice Set 2	5 responses representing the range of responses for the Knowledge and Use of Language Conventions trait	10 responses representing the range of responses for the conventions trait	Yes
Practice Set 3	10 responses representing the range of responses for both traits appropriate to the task type		Yes
Practice Set 4	10 responses representing the range of responses for both traits appropriate to the task type		Yes
Qualifying Set 1	10 responses comparable to the anchor set responses (included both traits appropriate to the task type)		No
Qualifying Set 2	10 responses comparable to the anchor set responses (included both traits appropriate to the task type)		No
Qualifying Set 3	10 responses comparable to the anchor set responses (included both traits appropriate to the task type)		No
Direct Copy Set**	3-5 responses composed entirely or partially of text copied from passage or passages (included both traits appropriate to the task type)	3-5 responses composed entirely or partially of text copied from passage or passages (included both traits appropriate to the task type)	Yes

**For the ELA Knowledge and Use of Language Conventions trait, there were two mixed-prompt anchor sets per grade level (one for the narrative task and the other for the literary analysis and research simulation tasks). In addition to the mixed-prompt anchor set, depending on the task, the practice sets for prototype and abbreviated items required readers to practice scoring the Knowledge and Use of Language Conventions trait along with the Reading Comprehension and Written Expression trait (for LAT and RST items) or with the Written Expression trait (NWT). Readers were also required to qualify on the Knowledge and Use of Language Conventions trait during each prototype item qualifying session.*

***These PARCC-approved sets provided additional annotated sample responses explaining the scoring rationale for responses composed entirely or partially of text copied from the source passage(s) associated with an item. DRC scoring supervisors reviewed these item-specific sets with the readers prior to scoring the associated item.*

Some items selected were previously only field-tested by PARCC. Consequently, the abbreviated training materials available for use with these items were abridged versions of typical abbreviated sets of materials. They consisted of:

- An Anchor Set (for ELA, some have annotations and some lack examples of the top scores)
- One Practice Set of 5 responses (scored but not annotated in the case of ELA)
- Approximately 10 validity responses

Since these materials were somewhat limited compared to typical abbreviated materials (the main difference being a lack of formal written annotations and fewer practice responses), DRC bolstered the training by using the PARCC-approved field test validity responses provided by New Meridian as additional practice responses. This work was done during the first administration of these items in 2019 and these augmented, LDOE-approved materials were used in 2020-2021. It is important to note that readers still had to qualify via standard qualification procedures on the prototype items for all items by first going through full training with the appropriate prototype Anchor Set, Practice Sets 1-4, and Qualifying Sets 1-3 (as well as the Conventions sets).

Qualifying Standards

DRC followed the same qualification standards that Pearson used for PARCC and New Meridian. Scorers demonstrated their ability to apply the scoring criteria by qualifying (i.e., scoring with acceptable agreement with true scores on qualifying sets). After each qualifying set was scored, the DRC scoring director responsible for training led the scorers in a discussion of the set. Any scorer who did not qualify by the end of the qualifying process for an item was not allowed to score live student responses.

Table 5.4 Mathematics Qualifying Standards

	Perfect Agreement	Perfect Plus Adjacent Agreement
0, 1, 2 Rubric	80% on two of three sets	96% on two of three sets
0, 1, 2, 3 Rubric	70% on two of three sets	96% on two of three sets
0, 1, 2, 3, 4 Rubric	70% on two of three sets	95% on two of three sets

Table 5.5 Mathematics Qualifying Standards (Composite Items)*

Composite (multipart) Items	Perfect Agreement	Perfect Plus Adjacent Agreement
0, 1 Rubric	90% on two of three sets	100% on two of three sets
0, 1, 2 Rubric	80% on two of three sets	96% on two of three sets
0, 1, 2, 3 Rubric	70% on two of three sets	96% on two of three sets
0, 1, 2, 3, 4 Rubric	70% on two of three sets	95% on two of three sets

*For mathematics composite items, the appropriate qualifying standard had to be achieved on each part of the item. For example, if an item had Part A with a top score of 1, Part B with a top score of 2, and Part C with a top score of 3, a scorer/supervisor would need to achieve 90% perfect agreement on Part A, 80% perfect agreement on Part B, and 70% perfect agreement on Part C, with no more than one nonadjacent score per part across all three qualifying sets.

Table 5.6 ELA Qualifying Standards

Perfect Agreement	Perfect Plus Adjacent Agreement
70% average for both traits on two of three qualifying sets	96% across the three qualifying sets combined on both traits
70% on each trait at least once across three qualifying sets	

ELA readers were required to meet all three of the qualifications listed in Table 5.6. Perfect plus adjacent agreement of 96% means that out of the entire pool of scores that a reader gave across the three qualifying sets for an item, no more than 4% of those scores could be nonadjacent. In other words, no more than 2 of the 60 applied scores could be nonadjacent (3 sets x 10 responses/set x 2 traits = 60 applied scores).

Monitoring the Scoring Process

Standard 6.8 states:

> Those responsible for test scoring should establish scoring protocols. Test scoring that involves human judgment should include rubrics, procedures, and criteria for scoring. When scoring of complex responses is done by computer, the accuracy of the algorithm and processes should be documented (118).

The following section explains the monitoring procedures that DRC uses to ensure that handscoring evaluators follow established scoring criteria while items are being scored. Detailed scoring rubrics, which specify the criteria for scoring, are available for handscoring evaluators for all constructed-response items.

Reader Monitoring Procedures

Throughout the handscoring process, DRC project managers, scoring directors, and team leaders reviewed the statistics that were generated on a daily basis. DRC used one team leader for every 10 to 12 readers, which was the same ratio that Pearson used for PARCC and New Meridian. If scoring concerns were apparent among individual scorers, team leaders dealt with those issues on an individual basis. If a scorer appeared to need clarification of the scoring rules, DRC supervisors typically monitored one out of five of the scorer's readings, making adjustments to that ratio as needed. If a supervisor disagreed with a reader's scores during monitoring, they provided retraining in the form of direct feedback to the reader, using rubric language and applicable training responses.

Validity Sets and Inter-Rater Reliability

In addition to the feedback that supervisors provided to readers during regular read-behinds and the continuous monitoring of inter-rater reliability and score point distributions, DRC also conducted validity scoring. Validity responses were inserted among the live student responses.

The validity responses were added to DRC's image handscoring system prior to the beginning of scoring. Validity reports compared readers' scores to pre-determined scores and were used to help detect potential room drift and individual scorer drift. This data was used to make decisions regarding the retraining and/or release of scorers, as well as the rescoring of responses.

Approximately 10% of all live student responses were scored by a second reader to establish inter-rater reliability statistics for all constructed-response items. This procedure is called a "double-blind read" because the second reader does not know the first reader's score. DRC monitored inter-rater reliability based on the responses that were scored by two readers. If a scorer fell below the expected rate of agreement, the team leader or scoring director retrained the scorer. If a scorer failed to improve after retraining and feedback,

DRC removed the scorer from the project. In this situation, DRC removed all scores assigned by the scorer in question. The responses were then reassigned and rescored.

To monitor inter-rater reliability, DRC produced scoring summary reports on a daily basis. DRC's scoring summary reports display exact, adjacent, and nonadjacent agreement rates for each reader. These rates are calculated based on responses that are scored by two readers, and their definitions are included below.

- **Percentage Exact (%EX)**—total number of responses by reader where scores are the same, divided by the number of responses that were scored twice
- **Percentage Adjacent (%AD)**—total number of responses by reader where scores are one point apart, divided by the number of responses that were scored twice
- **Percentage Nonadjacent (%NA)**—total number of responses by reader where scores are more than one score point apart, divided by the number of responses that were scored twice

The following table provided by Pearson shows the expectations for validity and inter-rater reliability:

Table 5.7 Agreement Rate Requirements for Validity and Inter-Rater Reliability

Subject	Score Point Range	Perfect Agreement	Perfect Agreement + Adjacent
Mathematics	0–1	90%	100%
	0–2	80%	95%
	0–3	70%	95%
	0–4	65%	95%
ELA	Multi-trait 0–3 or 0–4 (varies by grade and trait)	65% (each trait)	96% each trait)

Each reader was required to maintain a level of exact agreement on validity responses and on inter-rater reliability as shown under "Perfect Agreement" in the table above. Additionally, readers were required to maintain an acceptably low rate of nonadjacent agreement. To monitor this, DRC summed each reader's exact and adjacent agreement rates and required each reader to maintain the levels shown under "Perfect Agreement + Adjacent" in the table above.

Calibration Sets

New Meridian provided DRC with approved calibration responses for all operational items that came from the New Meridian item pool. DRC pulled calibration responses for DRC-developed math items. DRC used these sets to perform calibration across the entire scorer population for an item if trends were detected (e.g., low agreement between certain score points if a certain type of response was missing from initial training). These calibrations were designed to help refocus scorers on how to properly use the scoring guidelines. They were selected to help illustrate particular points and familiarize scorers with the types of responses commonly seen during operational scoring. After readers scored a calibration set, the scoring director reviewed it with the readers, using rubric language and scoring concepts exemplified by the anchor responses to explain the reasoning behind each response's score.

Reports and Reader Feedback

Reader performance and intervention information were recorded in reader feedback logs. These logs tracked information about actions taken with individual readers to ensure scoring consistency in regard to reliability,

score point distribution, and validity performance. In addition to the reader feedback logs, DRC provided LDOE with handscoring quality control reports for review throughout the scoring window. Further detail about these reports can be found in Appendix C.

5.2 Inter-Rater Reliability

A minimum of 10% of the constructed responses in ELA and mathematics were scored independently by a second reader. This was the case regardless of whether the first reader was human or AI. The statistics for inter-rater reliability were calculated for all items at all grades. To determine the reliability of scoring, the percentage of perfect agreement and adjacent agreement between the first and second scores was examined.

A total of 79 operational items were scored by human readers across all LEAP 2025 high school ELA and mathematics assessments. The inter-rater reliability rates and the total numbers of reads are shown in Tables 5.8–5.10 for ELA items, Tables 5.11–5.13 for mathematics items, Tables 5.14–5.16 for Spanish mathematics items, and Table 5.17 for mathematics field test items.

As shown in Tables 5.8–5.10, raters demonstrated at least 99% perfect and adjacent agreement for all ELA handscored items. As shown in Tables 5.11–5.13, raters demonstrated at least 99% perfect and adjacent agreement for mathematics items. As shown in Tables 5.14–5.16, raters demonstrated 100% perfect and adjacent agreement for Spanish mathematics items. As shown in Table 5.17, raters demonstrated 100% perfect and adjacent agreement for mathematics field test items.

Table 5.8 Inter-Rater Agreement, English Language Arts Items, Fall 2020

Course	Task Type	Question	Trait	Total Reads	Read 2x	Inter-Rater Reliability %		
						EX	AD	EX + AD
English I	Research Simulation (AI)	902161	Reading Comprehension and Written Expression	≥7,920	≥1,770	84	16	100
			Knowledge and Use of Language Conventions	≥7,920	≥1,770	83	17	100
	Narrative Writing (AI)	906152	Written Expression	≥7,820	≥2,080	85	15	100
			Knowledge and Use of Language Conventions	≥7,820	≥2,080	88	12	100
English II	Research Simulation (AI)	902331	Reading Comprehension and Written Expression	≥9,960	≥2,270	89	11	100
			Knowledge and Use of Language Conventions	≥9,960	≥2,270	89	11	100
	Narrative Writing	902354	Written Expression	≥9,660	≥2,020	84	16	100
			Knowledge and Use of Language Conventions	≥9,660	≥2,020	82	18	100

Table 5.9 Inter-Rater Agreement, English Language Arts Items, Spring 2021

Course	Task Type	Question /Form	Trait	Total Reads	Read 2x	Inter-Rater Reliability %		
						EX	AD	EX + AD
English I	Narrative Writing (AI)	983215	Written Expression	≥52,330	≥11,680	86	14	100
			Knowledge and Use of Language Conventions	≥52,330	≥11,680	85	15	100
	Research Simulation (AI)	914552	Reading Comprehension and Written Expression	≥52,720	≥11,780	82	18	100
			Knowledge and Use of Language Conventions	≥52,720	≥11,780	80	20	100
English II	Narrative Writing (AI)	983642	Written Expression	≥45,570	≥10,630	83	17	100
			Knowledge and Use of Language Conventions	≥45,570	≥10,630	82	18	100
	Research Simulation	983688	Reading Comprehension and Written Expression	≥45,710	≥9,960	81	19	100
			Knowledge and Use of Language Conventions	≥45,710	≥9,960	80	20	100

Table 5.10 Inter-Rater Agreement, English Language Arts Items, Summer 2021

Course	Task Type	Question	Trait	Total Reads	Read 2x	Inter-Rater Reliability %		
						EX	AD	EX + AD
English I	Literary Analysis (AI)	902152	Reading Comprehension and Written Expression	≥3,010	≥1,040	94	6	100
			Knowledge and Use of Language Conventions	≥3,010	≥1,040	93	7	100
	Research Simulation (AI)	902161	Reading Comprehension and Written Expression	≥2,920	≥910	92	8	100
			Knowledge and Use of Language Conventions	≥2,920	≥910	89	11	100
English II	Research Simulation (AI)	902331	Reading Comprehension and Written Expression	≥2,960	≥880	90	10	100
			Knowledge and Use of Language Conventions	≥2,960	≥880	92	8	100
	Literary Analysis (AI)	906197	Reading Comprehension and Written Expression	≥2,950	≥850	93	7	100
			Knowledge and Use of Language Conventions	≥2,950	≥850	93	7	100

Table 5.11 Inter-Rater Agreement, Mathematics Items, Fall 2020

Course	Question	Part(s)**	Total Reads	Read 2x	Inter-Rater Reliability %		
					EX	AD	EX + AD
Algebra I	901832	Part A	≥5,650	≥1,030	100	0	100
		Part B	≥5,650	≥1,030	93	7	100
	938741	N/A	≥5,550	≥1,320	97	3	100
	980927	Part A	≥5,450	≥1,250	99	1	100
		Part B	≥5,450	≥1,250	97	3	100
		Part C	≥5,450	≥1,250	95	5	100
	938735	Part A	≥5,580	≥1,000	100	0	100
		Part B	≥5,580	≥1,000	100	0	100
	938744	N/A	≥5,430	≥1,320	99	1	100
	938737	N/A	≥5,390	≥1,580	98	2	100
	938769	N/A	≥5,350	≥1,440	98	2	100
Geometry	902012	N/A	≥6,060	≥1,310	97	3	100
	980937	N/A	≥5,950	≥1,320	99	1	100
	939083	N/A	≥5,940	≥1,340	97	3	100
	980942	Part A	≥6,010	≥1,240	95	5	100
		Part B	≥6,010	≥1,240	97	3	100
	939077	N/A	≥5,920	≥1,390	98	2	100
	980938	N/A	≥5,900	≥1,420	99	1	100
	980936	N/A	≥5,930	≥1,280	95	5	100

*Total Ex + AD does not add up to 100% due to rounding
**N/A if an item does not have multiple parts

Table 5.12 Inter-Rater Agreement, Mathematics Items, Spring 2021

Course	Question	Part(s)**	Total Reads	Read 2x	Inter-Rater Reliability %		
					EX	AD	EX + AD
Algebra I	980924	N/A	≥53,330	≥11,440	91	9	100
	980909	N/A	≥52,820	≥13,110	93	6	99* (na = 0)
	980927	Part A	≥52,960	≥11,280	99	1	100
		Part B	≥52,960	≥11,280	96	4	100
		Part C	≥52,960	≥11,280	93	7	100
	980911	Part A	≥51,480	≥11,770	97	3	100
		Part B	≥51,480	≥11,770	96	4	100
	901851	N/A	≥52,700	≥11,830	90	10	100
	938737	N/A	≥51,930	≥13,380	96	4	100
	980923	N/A	≥52,070	≥12,170	96	4	100
Geometry	902012	N/A	≥37,230	≥8,190	91	8	99
	980937	N/A	≥36,910	≥8,640	97	3	100
	980929	N/A	≥36,360	≥9,300	95	5	100
	902042	Part A	≥37,100	≥7,640	97	2	99* (na = 0)
		Part B	≥37,100	≥7,640	98	2	100
		Part C	≥37,100	≥7,640	96	2	98* (na = 1)
	980930	Part B	≥38,080	≥6,930	95	5	100
	980938	N/A	≥36,310	≥9,000	98	2	100
	980936	N/A	≥36,640	≥8,260	92	8	100

*Total Ex + AD does not add up to 100% due to rounding
**N/A if an item does not have multiple parts

Table 5.13 Inter-Rater Agreement, Mathematics Items, Summer 2021

Course	Question	Part(s)*	Total Reads	Read 2x	Inter-Rater Reliability %		
					EX	AD	EX + AD
Algebra I	901832	Part B	≥3,940	≥720	96	4	100
	938741	N/A	≥3,850	≥990	98	2	100
	980927	Part A	≥3,810	≥960	100	0	100
		Part B	≥3,810	≥960	100	0	100
		Part C	≥3,810	≥960	99	1	100
	938735	Part A	≥3,890	≥690	100	0	100
		Part B	≥3,890	≥690	99	1	100
	938744	N/A	≥3,820	≥1,000	99	1	100
	938737	N/A	≥3,780	≥1,120	98	2	100
	938769	N/A	≥3,740	≥1,110	99	1	100
Geometry	902012	N/A	≥1,080	≥170	98	2	100
	980937	N/A	≥1,090	≥190	100	0	100
	939083	N/A	≥1,060	≥150	100	0	100
	980942	Part A	≥1,100	≥210	99	1	100
		Part B	≥1,100	≥210	97	3	100
	939077	N/A	≥1,060	≥190	99	1	100
	980938	N/A	≥1,070	≥180	100	0	100
	980936	N/A	≥1,120	≥250	100	0	100

*N/A if an item does not have multiple parts

Table 5.14 Inter-Rater Agreement, Spanish Mathematics Items, Fall 2020

Course	Question	Part(s)*	Total Reads	Read 2x**	Inter-Rater Reliability %		
					EX	AD	EX + AD
Algebra I	901832	Part A	≥20	<10	NR	NR	NR
		Part B	≥20	<10	NR	NR	NR
	938741	N/A	≥20	<10	NR	NR	NR
	980927	Part A	≥20	<10	NR	NR	NR
		Part B	≥20	<10	NR	NR	NR
		Part C	≥20	<10	NR	NR	NR
	938735	Part A	≥20	<10	NR	NR	NR
		Part B	≥20	<10	NR	NR	NR
	938744	N/A	≥20	<10	NR	NR	NR
	938737	N/A	≥10	N/A	N/A	N/A	N/A
	938769	N/A	≥10	<10	NR	NR	NR
Geometry	902012	N/A	≥20	<10	NR	NR	NR
	980937	N/A	≥20	N/A	N/A	N/A	N/A
	939083	N/A	≥20	<10	NR	NR	NR
	980942	Part A	≥20	N/A	N/A	N/A	N/A
		Part B	≥20	N/A	N/A	N/A	N/A
	939077	N/A	≥20	N/A	N/A	N/A	N/A
	980938	N/A	≥20	<10	NR	NR	NR
	980936	N/A	≥20	N/A	N/A	N/A	N/A

*N/A if an item does not have multiple parts

** Due to low numbers of Spanish mathematics test takers in fall 2020, some Spanish mathematics responses were scored directly by expert scorers/supervisors and not routed for second reads. As a result, no inter-rater reliability percentages were generated for those items.

Table 5.15 Inter-Rater Agreement, Spanish Mathematics Items, Spring 2021

Course	Question	Part(s)*	Total Reads	Read 2x**	Inter-Rater Reliability %		
					EX	AD	EX + AD
Algebra I	980924	N/A	≥70	<10	NR	NR	NR
	980909	N/A	≥70	<10	NR	NR	NR
	980927	Part A	≥70	<10	NR	NR	NR
		Part B	≥70	<10	NR	NR	NR
		Part C	≥70	<10	NR	NR	NR
	980911	Part A	≥70	<10	NR	NR	NR
		Part B	≥70	<10	NR	NR	NR
	901851	N/A	≥70	N/A	N/A	N/A	N/A
	938737	N/A	≥70	<10	NR	NR	NR
	980923	N/A	≥70	N/A	N/A	N/A	N/A
Geometry	902012	N/A	≥40	N/A	N/A	N/A	N/A
	980937	N/A	≥30	N/A	N/A	N/A	N/A
	980929	N/A	≥30	N/A	N/A	N/A	N/A
	902042	Part A	≥40	N/A	N/A	N/A	N/A
		Part B	≥40	N/A	N/A	N/A	N/A
		Part C	≥40	N/A	N/A	N/A	N/A
	980930	Part B	≥40	<10	NR	NR	NR
	980938	N/A	≥30	<10	NR	NR	NR
	980936	N/A	≥40	N/A	N/A	N/A	N/A

*N/A if an item does not have multiple parts

** Second Reads may be less than 10% of Total Reads or N/A for some items, because the smaller quantities of responses allowed scoring to be done directly by expert scorers/supervisors or via paired scoring between a supervisor and scorer. As a result, fewer were routed through the 10% read-behind process.

Table 5.16 Inter-Rater Agreement, Spanish Mathematics Items, Summer 2021

Course	Question	Part(s)*	Total Reads	Read 2x**	Inter-Rater Reliability %		
					EX	AD	EX + AD
Algebra I	901832	Part B	≥20	<10	NR	NR	NR
	938741	N/A	≥10	<10	NR	NR	NR
	980927	Part A	≥10	<10	NR	NR	NR
		Part B	≥10	<10	NR	NR	NR
		Part C	≥10	<10	NR	NR	NR
	938735	Part A	≥20	<10	NR	NR	NR
		Part B	≥20	<10	NR	NR	NR
	938744	N/A	≥10	<10	NR	NR	NR
	938737	N/A	≥10	<10	NR	NR	NR
	938769	N/A	≥10	<10	NR	NR	NR
Geometry	902012	N/A	<10	<10	NR	NR	NR
	980937	N/A	<10	N/A	N/A	N/A	N/A
	939083	N/A	<10	N/A	N/A	N/A	N/A
	980942	Part A	<10	N/A	N/A	N/A	N/A
		Part B	<10	N/A	N/A	N/A	N/A
	939077	N/A	<10	N/A	N/A	N/A	N/A
	980938	N/A	<10	N/A	N/A	N/A	N/A
	980936	N/A	<10	N/A	N/A	N/A	N/A

*N/A if an item does not have multiple parts
** Due to low numbers of Spanish mathematics test takers in spring 2021, some Spanish mathematics responses were scored directly by expert scorers/supervisors and not routed for second reads. As a result, no inter-rater reliability percentages were generated for those items.

5.3 Technology-Enhanced Item Scoring Process

All technology-enhanced items, as well as EBSR, MPSR, and SA items, were processed through DRC's autoscoring engine and scored according to the assigned scoring rules established during content development by PARCC or DRC in conjunction with LDOE. DRC ensured that all rubrics and scoring rules were verified for accuracy before scoring any technology-enhanced items. DRC established an adjudication process for technology-enhanced items and short answer responses to verify that correct answers were identified. DRC's technology-enhanced scoring process included the following procedures:

- A scoring rubric was created for each technology-enhanced item. The rubric described the one and only correct answer for dichotomously scored items (i.e., items scored as either right or wrong). If partial credit was possible, the rubric described in detail the type of response that could receive credit for each score point.
- The information from the scoring rubric was entered into the scoring system within the item banking system so that the rubric resided in one place along with the item image and other metadata. This scoring information included details that varied by item type. For example, for a drag-and-drop item, the information included which object is to be placed in each drop region to receive credit.
- The information was then verified by another autoscoring expert.
- After testing started, reports were generated that showed every response, how many students gave that response, and the score the scoring system provided for that response.
- The scoring was then checked against the scoring rubric using two levels of verification.
- If any discrepancies were found, the scoring information was modified and verified again. The scoring process was then rerun. This checking and modification process continued until no other issues were found.
- As a final check, a report was generated that showed all student responses, their frequencies, and their received scores.

5.4 Multiple-Choice and Multiple-Select Item Scoring Process

Responses to multiple-choice and multiple-select items were captured during test administration. In the case of braille forms, student responses to these items were transcribed into the online system by a test administrator.

5.5 Summary

The information presented in this chapter summarizes the scoring procedures for different types of items and the steps taken by DRC to ensure accuracy in the autoscoring and handscoring processes. The inter-rater reliability statistics presented in Section 5.2 demonstrate that the items were scored reliably. These efforts by DRC address multiple best practices of the testing industry but are particularly related to AERA, APA, & NCME (2014) Standards 4.18, 4.20, 6.8, and 6.9:

Standard 4.18 Procedures for scoring and, if relevant, scoring criteria, should be presented by the test developer with sufficient detail and clarity to maximize the accuracy of scoring. Instructions for using rating

scales or for deriving scores obtained by coding, scaling, or classifying constructed responses should be clear. This is especially critical for extended-response items such as performance tasks, portfolios, and essays (91).

Standard 4.20 The process for selecting, training, qualifying, and monitoring scorers should be specified by the test developer. The training materials, such as the scoring rubrics and examples of test takers' responses that illustrate the levels on the rubric score scale, and the procedures for training scorers should result in a degree of accuracy and agreement among scorers that allows the scores to be interpreted as originally intended by the test developer. Specifications should also describe processes for assessing scorer consistency and potential drift over time in raters' scoring (92).

Standard 6.8 Those responsible for test scoring should establish scoring protocols. Test scoring that involves human judgment should include rubrics, procedures, and criteria for scoring. When scoring of complex responses is done by computer, the accuracy of the algorithm and processes should be documented (118).

Standard 6.9 Those responsible for test scoring should establish and document quality control processes and criteria. Adequate training should be provided. The quality of scoring should be monitored and documented. Any systematic source of scoring errors should be documented and corrected (118).

Chapter 6: Operational Data Analyses

This chapter of the LEAP 2025 High School technical report describes the analyses that were conducted on the operational data. These include a classical item analysis and examination of the raw scores and an item response theory (IRT) analysis involving calibrating, scaling, and linking.

This section presents the classical item statistics, including aggregate raw score statistics and individual item-level statistics. Next, this section discusses the IRT models used for calibrating the data and addresses the purpose of data calibration and scaling for each content area. The calibration samples are then presented, followed by the data calibration results, including the model-data fit for the Louisiana student data. If the IRT models fit the empirical item response distributions for the population about which generalizations are to be made (i.e., Louisiana students), then the claim that the scores are valid indicators of an underlying ability is strengthened. The lowest obtainable scale score (LOSS) and highest obtainable scale score (HOSS) for the LEAP 2025 tests are also presented.

Chapter 6 demonstrates how LEAP 2025 assessments adhere to American Educational Research Association, American Psychological Association, & National Council on Measurement in Education (AERA, APA, & NCME, 2014) Standards 1.8, 4.14, 5.2, 5.13, 5.15, and 7.2. Each standard is explicated within the appropriate section of this chapter. Standard 7.2 provides general guidance that is relevant to this chapter. It states the following:

> The population for whom a test is intended and specifications for the test should be documented (126).

For all 2019-2021 LEAP 2025 high school analyses, the Louisiana student population was used. In Section 6.3, the characteristics of calibration samples, such as subgroups, are discussed. Chapter 3 presents the test specifications. Information regarding reported data is discussed in detail in Chapter 7.

6.1 Classical Item Statistics

In this section, summary test statistics for each form and subject area of the LEAP 2025 high school tests are presented. These statistics are followed by item-level statistics for each subject area of the LEAP 2025 test. These statistics were produced using census data with first-time test takers. Students whose results were included in the item-level statistics summary needed to meet at least the following psychometric analysis criteria (note that the criteria used to filter data for item statistics analyses are slightly different than those used to produce students' performance statistics in this report):

- Student has total raw score in the data
- Student did not take administration error form
- Student did not take braille form
- Student did not take Spanish form
- Student's test score was not voided
- Student took the assessment for the first time (initial testers)
- Student finished all sessions
- Student's constructed-response items were scored

Test-Level Statistics

Table 6.1 presents the number of items, score points, mean and standard deviation of the raw scores, and the average form difficulty for each subject for each administration. Form difficulty for a student was calculated by dividing the student's raw score by the total score points of the test.

As can be seen in the table, average form difficulty was similar in the fall and spring administrations. Average form difficulty in the summer administration was lower for all tests than in the fall and spring administrations, likely due to the fact that the summer form was a retest form. The average form difficulty for ELA ranged from 0.42 to 0.46 with the fall and spring administrations. The difficulty of the spring administration forms was 0.45 English I and 0.46 for English II. The average form difficulty for mathematics ranged from 0.32 to 0.33 for the fall and spring administrations. The average form difficulty of the spring administration of mathematics was 0.33 for Algebra I and 0.33 for Geometry. In general, the 2021 LEAP 2025 High School tests were relatively difficult, and the mathematics tests were more difficult than the ELA tests.

Table 6.1 LEAP 2025 High School Means and Standard Deviations for Raw Scores and Form Difficulty

Administration	Course	Form	Total Items*	Total Points	Mean Raw Score (Std. Dev.)	Average Form Difficulty (Std. Dev.)
Fall 2020	English I	A	33	90	36.03 (17.76)	0.42 (0.12)
	English II	A	33	90	38.90 (17.25)	0.46 (0.15)
	Algebra I	D	39	68	18.81 (10.02)	0.32 (0.15)
	Geometry	D	39	68	18.83 (12.01)	0.33 (0.14)
Spring 2021	English I	E	33	90	36.71 (17.95)	0.45 (0.11)
	English II	E	33	90	39.92 (18.19)	0.46 (0.11)
	Algebra I	E	39	68	19.40 (11.54)	0.33 (0.15)
	Geometry	E	39	68	19.16 (12.03)	0.33 (0.14)
Summer 2021	English I	B	34	94	16.48 (9.75)	0.22 (0.11)
	English II	B	34	94	18.88 (10.01)	0.24 (0.11)
	Algebra I	D	39	68	11.97 (6.01)	0.22 (0.13)
	Geometry	D	39	68	9.96 (5.03)	0.18 (0.13)

*For English I and English II, each writing prompt component is counted as one item. The WE writing component is weighted in total points.

Table 6.2 presents the number of items, mean and standard deviation of the item *p*-values, and item-total correlations (i.e., item discrimination values) for each subject for each administration.

The mean *p*-value is the average of all item *p*-values in a specific subject area and administration. The mean item-total correlation (R_{it}) is the average of all item point-biserial correlations of a specific subject area. The *p*-value and item-total correlation are explained in the next section.

Table 6.2 LEAP 2025 High School *p*-Values and Item-Total Correlation (R_{it}) Descriptive Statistics

				Item *p*-Value				Average Item-Total Correlation			
Admin.	**Course**	**Form**	**Total Items***	**Mean**	**Std. Dev.**	**Min.**	**Max**	**Mean**	**Std. Dev.**	**Min.**	**Max**
Fall 2020	English I	A	33	0.42	0.12	0.09	0.73	0.46	0.18	0.16	0.81
	English II	A	33	0.46	0.15	0.23	0.81	0.47	0.17	0.27	0.82
	Algebra I	D	39	0.32	0.15	0.09	0.80	0.36	0.15	0.12	0.61
	Geometry	D	39	0.33	0.14	0.06	0.65	0.46	0.17	0.14	0.75
Spring 2021	English I	E	33	0.45	0.11	0.27	0.78	0.48	0.15	0.21	0.80
	English II	E	33	0.46	0.11	0.28	0.69	0.47	0.16	0.25	0.83
	Algebra I	E	39	0.33	0.15	0.09	0.81	0.41	0.16	0.14	0.69
	Geometry	E	39	0.33	0.14	0.06	0.61	0.45	0.15	0.22	0.75
Summer 2021	English I	B	34	0.22	0.11	0.05	0.46	0.32	0.18	-0.08	0.67
	English II	B	34	0.24	0.11	0.10	0.54	0.31	0.20	0.02	0.72
	Algebra I	D	39	0.22	0.13	0.02	0.61	0.26	0.16	0.04	0.62
	Geometry	D	39	0.18	0.13	0.01	0.45	0.26	0.16	0.00	0.54

*For English I and English II, each writing prompt component is counted as one item. The WE writing component is weighted in total points.

Item-Level Statistics

Tables in Appendix D present the item statistics for each operational item included in the regular forms, organized by content area and administration. The tables include item number, *p*-value, item-total correlation (R_{it}), omit rates, total N, adjusted N (adjusted N excludes omitted responses, responses that were not scored, or responses that received a non-score code), the percentage at each score point for polytomous items, and the percentage that chose each option for multiple-choice (MC) items. The *p*-value and item-total correlations calculations used the adjusted N to determine the values. The rest of the statistics in the table are based on the total N.

The summer administration population is not state representative, and the number of students was very small, so the interpretation of statistics in the summer administration should be done with caution.

p-Value

The *p*-value is a measure of item difficulty. For an MC item, the *p*-value is calculated by dividing the number of students who correctly responded to an item by the total number of students who attempted the item. The value is reported as a proportion. For a non-MC item, the *p*-value is calculated by dividing the average score for the item by the maximum points possible. This value is also reported as a proportion.

In terms of *p*-values, test scores tend to be more precise when their average *p*-values are between the mid-0.50s and the low 0.70s. However, it is important to select items based on content rather than on purely statistical criteria when building a criterion-referenced test. As shown in Table 6.2, the average *p*-values of the fall and spring administrations ranged from 0.33 to 0.46. The range of average *p*-values was lower in the summer administration, ranging from 0.18 to 0.24. The average *p*-values of the English I and English II forms were higher than the average p-values of the Algebra I forms.

It is important that one examines the range of *p*-values, not just the average *p*-value, to determine whether a test measures well. It is desirable for a test to measure well throughout the range of skills present in the test

form. That is, it is important that the items measure the performance of students at all levels of achievement, not just students in the center of the distribution. Having a range of *p*-values also helps to prevent floor and/or ceiling effects so that the test does not have large numbers of students at the minimum or maximum possible scores. The fall and spring English forms have items with *p*-values ranging from 0.09 to 0.78 (see Appendix D) and the summer English forms have items with p-values ranging from 0.05 to 0.54. The *p*-values on the mathematics forms range from 0.06 to 0.81 (see Appendix D) for the fall and spring administrations and from 0.01 to 0.61 for the summer administration. Such a broad range of *p*-values, which indicates that the items measure well throughout the range of skill levels at a given grade, supports the accuracy of the LEAP 2025 high school test scores.

Item-Total Correlations

An item-total correlation is the correlation between an item score and the total test score, where the item score is not included in the total score. It indicates how well an item differentiates student performance across all levels of achievement. In general, items with correlations below 0.20 are said to be poorly discriminating. The majority of the items on the LEAP 2025 High School forms had item-total correlations above this threshold. Any item with an item-total correlation below the 0.20 threshold was further analyzed to ensure that the item was correctly keyed. It was not unusual for items to have lower item-total correlations from the summer administration due to being administered to a re-test population.

Omit Rates

The omit rate for each item indicates the percentage of students who did not answer the item. Omit rates can be used to examine possible speededness issues on tests. A test may be speeded if students do not have adequate time to answer all questions on the test. In general, an item is said to have a high omit rate if more than 5% of students failed to respond to the item. Evidence of speededness is considered a threat to validity because student test scores may not reflect their ability. Additionally, content validity may be threatened because the items that were not completed are needed to fulfill content blueprint specifications (Lu & Sireci, 2007).

This examination of omit rates complies with Standard 4.14 of the *Standards*. This standard is concerned with the speededness of a test and states the following:

> For a test that has a time limit, test development research should examine the degree to which scores include a speed component and should evaluate the appropriateness of that component, given the domain the test is designed to measure (90).

The results in this section will show that, overall, student test scores are not adversely affected by the rate at which the students complete the test. In general, students have ample time to complete all sections of the test, and there is not a threat to construct or content validity.

The results presented in the Tables in Appendix D show that the percentage of students who omitted most of the items on the fall and spring LEAP 2025 High School tests was less than 5, suggesting that most students were able to complete the test in the prescribed amount of time. There were a small number of Algebra I and Geometry items that exceeded the omit rate of 5%. This is likely due to the difficulty of the items, given that these items also have low *p*-values. Lu & Sireci (2007) report that the Education Testing Service has used an approach where a test was considered unspeeded if at least 80% of examinees reach the last item and all examinees reach at least 75% of the items. The reported omit rates fall within these ranges.

These item level statistics are reviewed at the beginning of the operational analysis process to ensure that items are not flawed, and a careful review is given to determine that the answer key is correct.

An MC item is reviewed during the key check process if

- it has a *p*-value less than 0.25 or more than .95,
- a greater number of high-performing students (top 20%) are choosing a distractor than are choosing the key,
- the item-total correlation is less than 0.20,
- any of the incorrect answer options yields a positive distractor-total correlation, or
- the percentage of students omitting or not reaching each item is 5 or greater.

Other types of autoscored items are also flagged during the key check for review if the

- *p*-value is less than 0.30 or more than .80,
- percentage of students who reached any possible score point is less than 3%,
- item-total correlation is less than 0.30, or
- percentage of students omitting or not reaching the item is 15% or greater.

Item Response Theory (IRT)

Item parameters for items included in the LEAP 2025 High School tests were estimated using a marginal maximum-likelihood (MML) procedure and the 2-parameter logistic (2PL) model for MC items and the generalized partial credit (GPC) model (Muraki, 1992) for non-MC items. Under the 2PL model, the probability that a student with a trait or scale score of θ will respond correctly to MC item *j* is

$$P_j(\theta) = \frac{1}{1+\exp\left[-Da_j(\theta-b_j)\right]},$$

where *D* is 1.7, a_j is the item discrimination, and b_j is the item difficulty. Under the GPC model, the probability that a student with a trait or scale score of θ will respond in category *x* to partial-credit item *j* is

$$P_{jx}(\theta) = \frac{exp[\sum_{k=0}^{x} Da_j(\theta-b_j+d_{jk})]}{\sum_{h=0}^{m_i} exp[\sum_{k=0}^{h} Da_j(\theta-b_j+d_{jk})]},$$

where d_{jk} is the relative difficulty of score category *x* of item *j*, and m_i is the maximum item score for item *j*.

The software IRTPRO (Cai, Thissen, & du Toit, 2011) was used for the IRT calibrations. IRTPRO is a multipurpose program that implements a variety of IRT models associated with mixed-item formats and associated statistics. IRTPRO has been used to calibrate large data sets, such as those of PARCC and Smarter Balanced assessments. The program implements MML estimation techniques for items and MLE estimation of theta.

6.2 Calibration Sample

Item calibration and linking for the LEAP 2025 high school assessments were not performed in the spring of 2021. The forms used in the 2020-2021 administration were intact forms previously used in the 2018-2019 administration. For information regarding calibration and linking of these forms, please see the *2018-2019 LEAP 2025 High School Operational Technical Report* Comparability: Form Equating.

6.3 Calibration and Linking

The primary purpose of form equating is to establish score equivalency between two (or more) forms. Equivalency is established by first building the forms to be equated according to tight content specifications. Then the form scores are placed on the same scale (by equating), such that students performing on two scaled assessments at the same level of underlying achievement should receive the same scale score on both forms, although they may not receive the same number-correct score (or raw score). The raw-to-scale-score

relationship performs this leveling function based on form-equating studies. Theoretically, differences in the raw-to-scale-score relationship between the two forms can be partially due to differences in the samples utilized for calibration and differences in item difficulty. LDOE and DRC strive to maintain equivalent samples or use near-census samples over the years, minimizing the potential differences caused by the different samples. Differences in the raw-to-scale-score relationship, therefore, can be primarily attributed to the differences in item difficulty.

In the spring of 2021, the forms used were intact and when originally administered in 2019, they were post-equated and linked to the LEAP 2025 scale. The equating was conducted using the test characteristic transformation function method in the common-item non-equivalent-groups design (Stocking & Lord, 1983). The fall 2020 and summer 2021 forms were also intact forms.

Table 6.13 through Table 6.16 provide scale scores at selected percentiles that can be used to compare the distributional characteristics of the LEAP 2025 2020–2021 forms to previous administrations. Although these scale scores are rounded values, there were differences in the scale score values for a given percentile across the forms. These variations could arise for several reasons: (1) differences in the proficiency (i.e., achievement) of the students in the samples or growth in student achievement across years; (2) unevenness in the respective distributions that combine with the number-correct-to-scale-score scoring method, leaving “gaps” in the scale; or (3) other sources of equating error. Other sources of equating error can include subtle content differences between forms, handscoring differences, or unusual student samples. Some equating errors will always be present between forms. This means that the forms will not measure identically, even under optimal testing conditions. In general, however, the test characteristic function equating techniques will “level” the equated forms through the raw-to-scale-score adjustment.

Table 6.3 Comparisons of Scale Scores at Selected Percentiles—English I

	2017	2018	2019	2020	2018	2019	2019	2021	2018	2019	2020	2021
Percentile	Fall	Fall	Fall	Fall	Spring	Spring Form D	Spring Form E	Spring Form E	Summer	Summer	Summer	Summer
99	818	821	821	821	824	820	824	821	756	750	753	757
95	795	796	795	795	799	796	802	797	736	734	732	738
90	784	784	782	782	788	785	790	786	727	725	719	728
85	776	774	772	772	778	777	782	778	721	719	713	720
80	771	767	765	763	773	769	776	770	715	715	709	714
75	765	760	756	758	767	764	770	765	710	713	704	710
70	762	755	749	752	762	759	765	759	708	709	701	707
65	757	749	743	747	757	753	759	754	704	707	696	703
60	753	742	736	740	754	748	756	749	701	704	696	700
55	750	737	729	735	749	743	753	744	699	702	693	697
50	745	731	724	730	745	738	748	741	693	699	690	694
45	742	726	718	725	740	732	744	736	691	696	686	691
40	737	720	712	719	737	729	739	731	688	693	683	688
35	734	712	705	713	731	723	734	726	684	693	683	685
30	728	707	700	709	728	717	729	720	684	689	679	685
25	723	700	694	704	722	713	726	714	681	686	679	681
20	717	694	691	699	716	706	718	707	677	682	674	676
15	711	688	684	693	707	699	712	700	673	677	669	671
10	702	681	679	686	697	686	705	691	669	677	657	666
5	693	666	669	677	685	674	691	679	658	665	650	660
1	671	650	650	657	660	650	669	656	650	650	650	650

Table 6.4 Comparisons of Scale Scores at Selected Percentiles—English II

	2017	2018	2019	2020	2018	2019	2019	2021	2018	2019	2020	2021
Percentile	Fall	Fall	Fall	Fall	Spring	Spring Form D	Spring Form E	Spring Form E	Summer	Summer	Summer	Summer
99	838	846	850	842	846	842	847	850	760	761	756	773
95	805	811	818	811	817	810	818	821	737	733	737	740
90	784	788	799	793	799	795	802	805	725	722	719	726
85	772	775	786	782	788	787	791	794	720	714	705	719
80	763	763	778	772	780	779	783	783	715	709	700	714
75	754	754	768	765	773	771	778	775	709	704	691	709
70	748	747	758	758	765	764	770	768	703	702	688	707
65	740	738	749	752	761	759	765	763	700	699	684	702
60	734	731	743	746	754	752	761	756	696	693	677	699
55	726	724	734	740	749	745	754	749	689	690	672	693
50	720	717	728	736	745	741	749	745	685	687	672	690
45	714	712	722	730	738	734	743	739	681	684	668	687
40	707	707	713	724	733	728	739	732	676	680	662	684
35	702	702	706	720	726	722	732	726	676	676	656	680
30	694	693	698	714	722	715	726	720	671	672	650	672
25	688	687	693	707	714	705	720	711	666	668	650	668
20	684	680	683	699	707	697	713	703	659	663	650	663
15	677	672	675	691	699	688	703	693	659	663	650	658
10	668	663	667	681	687	672	693	679	652	652	650	652
5	658	652	656	668	668	656	675	667	650	650	650	650
1	650	650	650	650	650	650	650	650	650	650	650	650

Table 6.5 Comparisons of Scale Scores at Selected Percentiles—Algebra I

	2017	2018	2019	2020	2018	2019	2019	2021	2018	2019	2020	2021
Percentile	Fall	Fall	Fall	Fall	Spring	Spring Form D	Spring Form E	Spring Form E	Summer	Summer	Summer	Summer
99	801	810	811	808	827	836	839	822	758	759	819	786
95	778	785	788	782	800	799	803	795	738	739	751	742
90	761	772	774	770	787	786	789	780	730	732	739	732
85	749	760	766	763	777	776	780	770	724	728	728	728
80	743	754	759	753	769	768	772	761	721	725	725	721
75	733	744	751	748	763	761	766	756	718	721	721	721
70	727	738	742	742	757	753	761	749	714	717	717	717
65	724	734	735	739	751	748	756	743	714	717	717	712
60	721	727	732	732	748	745	751	737	710	713	712	712
55	718	723	725	728	744	739	746	734	710	713	708	708
50	714	719	721	725	738	735	740	730	705	708	708	708
45	710	715	717	721	734	728	737	727	705	708	703	708
40	710	711	712	717	731	725	734	723	705	704	703	703
35	705	711	712	712	727	721	730	719	700	704	697	703
30	705	707	708	708	723	717	723	714	700	698	697	697
25	700	702	703	708	715	712	719	710	695	698	697	697
20	695	696	697	703	711	708	714	704	695	692	691	691
15	695	696	697	697	707	703	710	699	688	692	685	691
10	688	690	691	691	702	697	704	692	688	685	685	685
5	680	683	677	685	690	691	692	685	680	675	677	677
1	669	650	668	668	673	668	677	668	655	650	656	656

Table 6.6 Comparisons of Scale Scores at Selected Percentiles—Geometry

	2017	2018	2019	2020	2018	2019	2019	2021	2018	2019	2020	2021
Percentile	**Fall**	**Fall**	**Fall**	**Fall**	**Spring**	**Spring Form D**	**Spring Form E**	**Spring Form E**	**Summer**	**Summer**	**Summer**	**Summer**
99	797	799	809	801	796	801	801	801	820	816	758	764
95	779	781	788	780	779	783	784	784	760	785	744	736
90	768	771	777	769	771	774	774	774	724	738	734	729
85	761	764	771	762	764	767	768	768	719	723	729	726
80	755	760	764	756	758	761	763	763	716	720	726	723
75	749	754	759	752	754	755	758	758	712	717	719	719
70	744	751	755	746	749	750	753	753	712	717	716	719
65	740	746	750	742	746	746	750	750	709	714	716	716
60	736	742	746	738	742	742	746	746	709	710	716	716
55	732	736	742	734	738	738	742	742	705	710	716	716
50	727	731	738	731	734	734	738	738	705	706	712	712
45	724	729	734	726	731	731	733	733	705	706	712	712
40	722	724	731	723	727	729	731	731	701	701	712	707
35	716	721	729	719	724	726	728	728	701	701	707	707
30	712	715	726	719	721	723	725	725	696	696	707	707
25	709	711	719	716	718	719	722	722	696	696	707	701
20	705	707	716	712	715	716	719	719	696	690	701	701
15	701	702	712	707	707	712	711	711	691	690	694	694
10	696	697	707	701	702	707	706	706	685	684	694	694
5	691	692	701	694	697	694	701	701	678	676	686	686
1	678	677	675	675	677	675	686	686	670	666	661	661

Additional evidence of comparability can be found by reviewing the test characteristic curves (TCCs) across administrations of the LEAP 2025 assessments, as can be seen in Figure 6.12. Note that the spring 2021 form was previously administered in spring 2019; therefore, they have the same TCC. For most content areas, the TCCs across years were similar across ability ranges.

Figure 6.13 shows SEMs for the across administration years for the LEAP 2025 HS assessments. For most content areas, the SEMs were similar across ability ranges, especially in the middle ability ranges.

Figure 6.1 TCCs Across Years: Spring Administrations

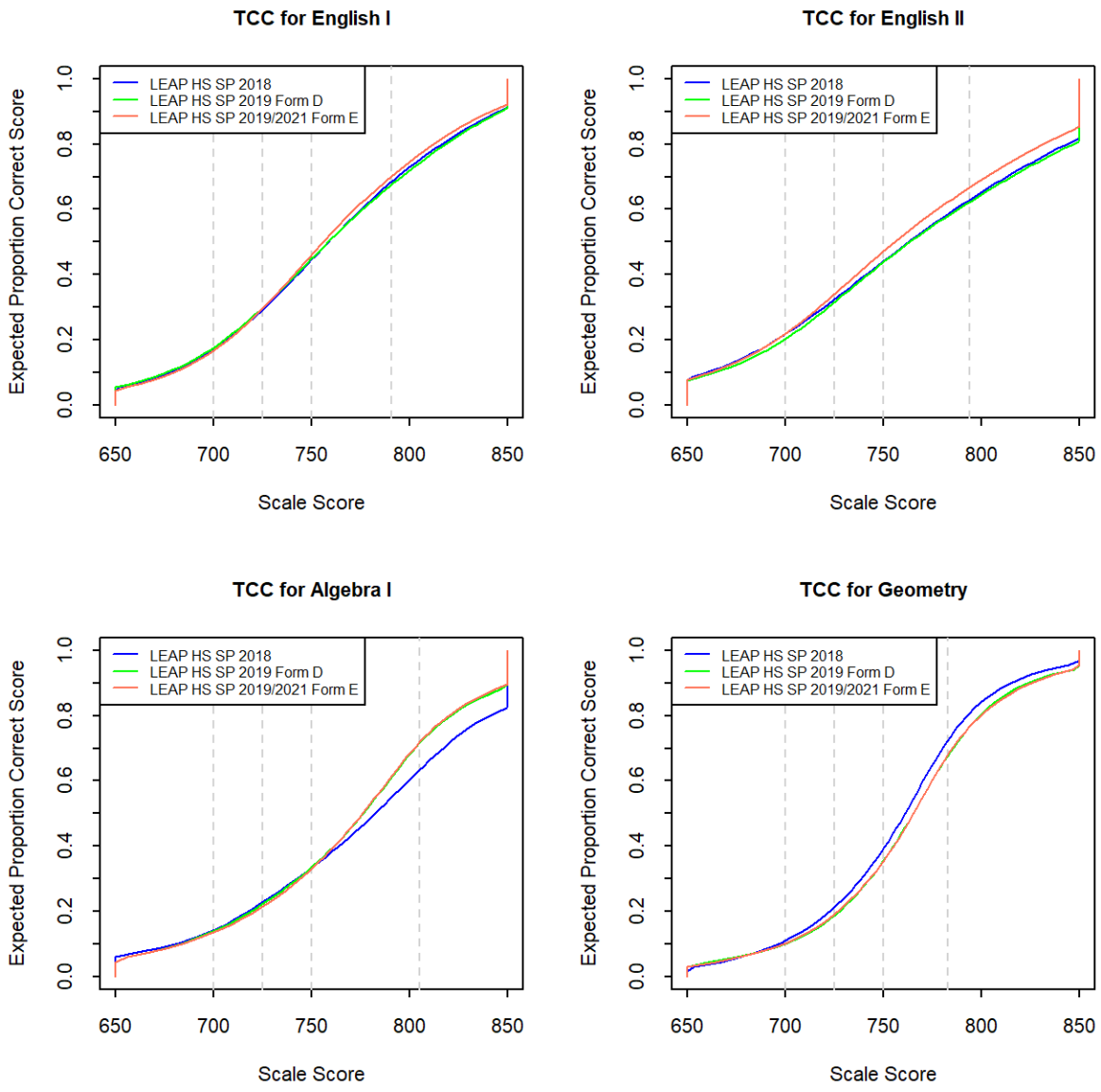
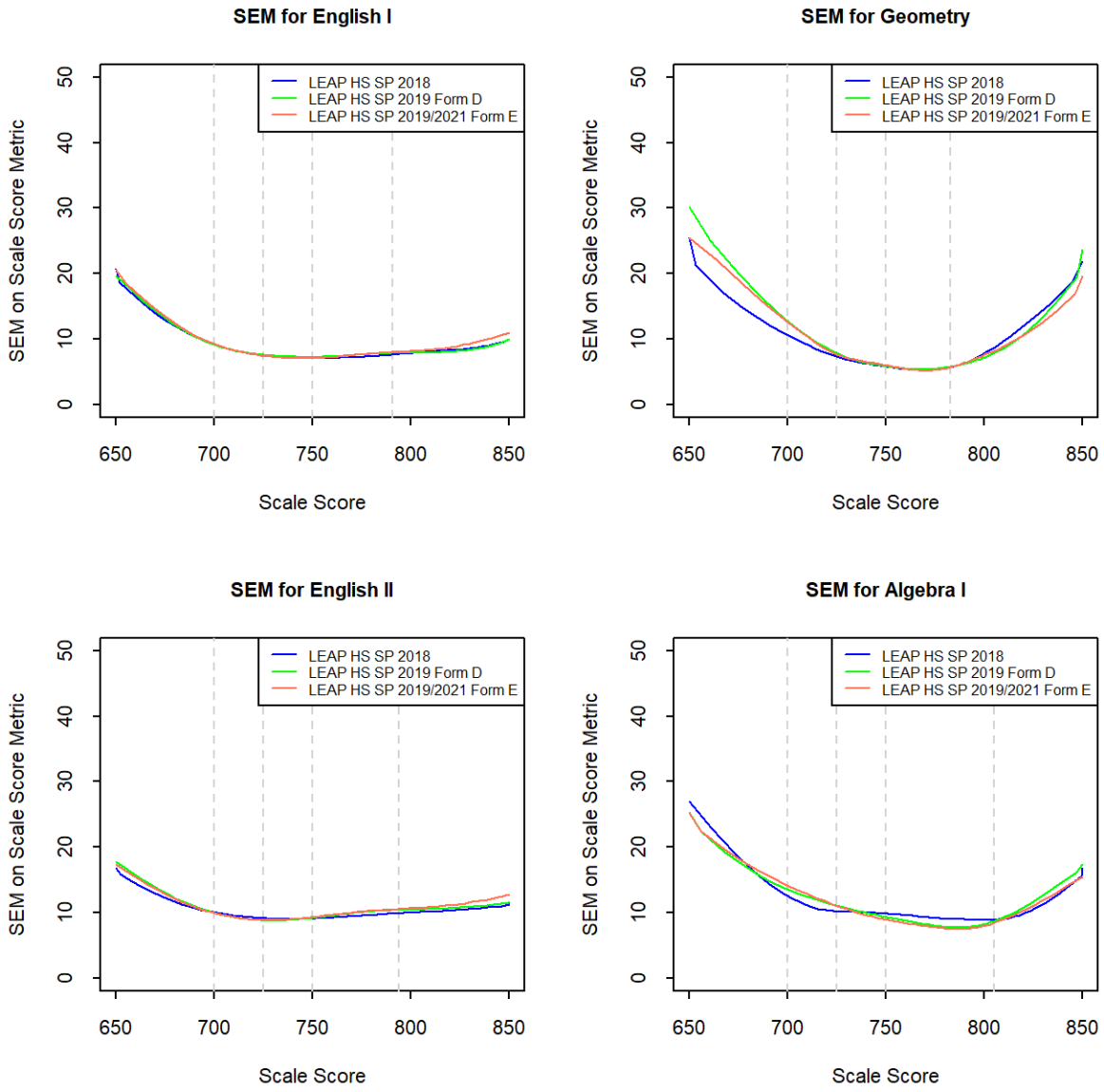


Figure 6.2 SEM Across Years: Spring Administrations



6.4 Summary

In summary, the overall purpose of the operational data analyses is to ensure that the test items, as well as the overall test, are functioning appropriately. Operational data analyses also help maintain the test scale so that test results may be appropriately compared across years. The data analyses undertaken by DRC address multiple best practices of the testing industry but are particularly related to the following standards:

> *Standard 1.8* The composition of any sample of test takers from which validity evidence is obtained should be described in as much detail as is practical and permissible, including major relevant socio-demographic and developmental characteristics (25).

Standard 4.14 For a test that has a time limit, test development research should examine the degree to which scores include a speed component and should evaluate the appropriateness of that component, given the domain the test is designed to measure (90).

Standard 5.2 The procedures for constructing scales used for reporting scores and the rationale for these procedures should be described clearly (102).

Standard 5.13 When claims of form-to-form score equivalence are based on equating procedures, detailed technical information should be provided on the method by which equating functions were established and on the accuracy of the equating functions (105).

Standard 5.15 In equating studies that employ an anchor test design, the characteristics of the anchor test and its similarity to the forms being equated should be presented, including both content specifications and empirically determined relationships among test scores. If anchor items are used in the equating study, the representativeness and psychometric characteristics of the anchor items should be presented (105).

Standard 7.2 The population for whom a test is intended and specifications for the test should be documented. If normative data are provided, the procedures used to gather the data should be explained; the norming population should be described in terms of relevant demographic variables; and the year(s) in which the data were collected should be reported (126).

Chapter 7: Test Results

This chapter of the technical report contains information on the results of the Spring LEAP 2025 High School administration of English I, English II, Algebra I, and Geometry. The scale score results and achievement level information are presented here. Presenting the results by achievement level translates the quantitative scale provided through scale scores into a qualitative description of student achievement. The levels are *Advanced*, *Mastery*, *Basic*, *Approaching Basic*, and *Unsatisfactory*.

While the scale score provides an essential quantitative reference for student achievement, the achievement level information plainly outlines the meanings of the scores to parents, students, and educators. When combined, scale scores and achievement levels provide a comprehensive set of tools to assess Louisiana student achievement by course.

This chapter also provides descriptions of the score reports, data structure, and interpretive guide for the LEAP 2025 administrations. The American Educational Research Association, American Psychological Association, and National Council on Measurement in Education (AERA, APA, & NCME, 2014) *Standards for Educational & Psychological Testing* addressed in Chapter 7 are 5.1, 6.10, 7.0, and 12.18. Each standard is presented in the pertinent section of this chapter.

The results presented in this chapter are based on census data. The results presented here may differ slightly from the official state summary report of all student populations due to ongoing resolution of test materials and student information. The results in the tables in this chapter are presented as evidence of the reliability and validity of the scores from the LEAP 2025 high school ELA and mathematics assessments and should not be used for state accountability purposes.

7.1 Student Participation

The following are subgroups reported during the administration of the LEAP 2025 tests:

- Gender: Female and Male
- Race and Ethnicity: Hispanic/Latino, American Indian or Alaska Native, Asian, Black or African American, Native Hawaiian or Other Pacific Islander, White, and Two or More Races
- Education Classification
- Economic Status
- English Learner (EL)
- Migrant Status
- Homeless Status
- Military Affiliation
- Foster Care Status

The number of students who attempted each test, the number of students whose results were reportable from each test, and the number of students whose results were included in the technical report sample for each test are summarized by grade in Tables 7.1–7.4. The “Attempted” category includes all the students who attempted at least one item on the assessment. The “Reportable” category includes students who finished all sections in the assessment, which includes students in private school and home-study programs. The “Technical Report Sample” category represents the sample of students included in the analyses for this report, and they are the students who finished all sections of the assessment and counted toward the state total score; students in private school and home-study programs were excluded from this sample.

Table 7.1 Count of Students who Attempted, were Reportable, and Included in the Technical Report
Sample: English I

Administration	Group	Grade							
		6	7	8	9*	10	11	12	Total
Fall 2020	Attempted	<10	<10	≥10	≥5,000	≥870	≥790	≥500	≥7,190
	Reportable	<10	<10	≥10	≥4,960	≥850	≥760	≥480	≥7,080
	Technical Report	<10	<10	≥10	≥4,950	≥850	≥760	≥480	≥7,070
Spring 2021	Attempted	<10	<10	≥2,120	≥40,830	≥3,340	≥930	≥420	≥47,670
	Reportable	<10	<10	≥2,100	≥40,530	≥3,200	≥890	≥420	≥47,160
	Technical Report	<10	<10	≥2,050	≥40,160	≥3,200	≥880	≥420	≥46,730
Summer 2021	Attempted	<10	<10	≥10	≥1,970	≥430	≥210	≥10	≥2,640
	Reportable	<10	<10	≥10	≥1,940	≥420	≥200	≥10	≥2,600
	Technical Report	<10	<10	<10	≥1,930	≥420	≥200	≥10	≥2,580

* Grade 9 includes the grade that is coded as "T9."

Table 7.2 Count of Students who Attempted, were Reportable, and Included in the Technical Report
Sample: English II

Administration	Group	Grade							
		6	7	8	9*	10	11	12	Total
Fall 2020	Attempted	<10	<10	<10	≥1,130	≥6,350	≥890	≥630	≥9,010
	Reportable	<10	<10	<10	≥1,120	≥6,280	≥860	≥600	≥8,880
	Technical Report	<10	<10	<10	≥1,120	≥6,280	≥860	≥600	≥8,880
Spring 2021	Attempted	<10	<10	<10	≥2,230	≥35,950	≥2,590	≥700	≥41,480
	Reportable	<10	<10	<10	≥2,200	≥35,750	≥2,520	≥680	≥41,170
	Technical Report	<10	<10	<10	≥2,180	≥35,440	≥2,520	≥680	≥40,830
Summer 2021	Attempted	<10	<10	<10	≥120	≥2,170	≥360	≥10	≥2,670
	Reportable	<10	<10	<10	≥120	≥2,150	≥360	≥10	≥2,650
	Technical Report	<10	<10	<10	≥120	≥2,140	≥360	≥10	≥2,640

* Grade 9 includes the grade that is coded as "T9."

Table 7.3 Count of Students who Attempted, were Reportable, and included in the Technical Report Sample: Algebra I

Administration	Group	Grade							
		6	7	8	9*	10	11	12	Total
Fall 2020	Attempted	<10	<10	≥20	≥3,360	≥850	≥620	≥300	≥5,170
	Reportable	<10	<10	≥20	≥3,330	≥840	≥600	≥290	≥5,100
	Technical Report	<10	<10	≥20	≥3,330	≥840	≥600	≥290	≥5,100
Spring 2021	Attempted	<10	≥210	≥7,170	≥35,520	≥5,350	≥950	≥250	≥49,480
	Reportable	<10	≥210	≥7,130	≥35,220	≥5,190	≥920	≥250	≥48,950
	Technical Report	<10	≥210	≥7,020	≥34,890	≥5,180	≥920	≥250	≥48,500
Summer 2021	Attempted	<10	<10	≥100	≥2,630	≥690	≥160	≥20	≥3,610
	Reportable	<10	<10	≥100	≥2,590	≥680	≥150	≥20	≥3,560
	Technical Report	<10	<10	≥90	≥2,590	≥670	≥150	≥20	≥3,540

* Grade 9 includes the grade that is coded as "T9."

Table 7.4 Count of Students who Attempted, were Reportable, and Included in the Technical Report Sample: Geometry

Administration	Group	Grade							
		6	7	8	9*	10	11	12	Total
Fall 2020	Attempted	<10	<10	<10	≥1,220	≥3,030	≥1,320	≥170	≥5,750
	Reportable	<10	<10	<10	≥1,200	≥2,990	≥1,310	≥170	≥5,680
	Tech Report	<10	<10	<10	≥1,200	≥2,990	≥1,310	≥170	≥5,680
Spring 2021	Attempted	<10	≥20	≥240	≥6,210	≥23,710	≥4,280	≥320	≥34,810
	Reportable	<10	≥20	≥240	≥6,190	≥23,600	≥4,230	≥320	≥34,620
	Tech Report	<10	≥20	≥240	≥6,140	≥23,360	≥4,160	≥310	≥34,260
Summer 2021	Attempted	<10	<10	<10	≥50	≥770	≥230	<10	≥1,070
	Reportable	<10	<10	<10	≥50	≥770	≥230	<10	≥1,060
	Tech Report	<10	<10	<10	≥50	≥760	≥230	<10	≥1,050

* Grade 9 includes the grade that is coded as "T9."

The counts and percentages of students in demographic groups by grade for the group of students comprising the technical report sample for the spring 2021 administration are summarized in Table 7.5 through Table 7.12. The same information regarding the technical report samples for the fall 2018 and summer 2019 administrations can be found in Appendix E.

Table 7.5 Count of Students taking the Spring 2021 LEAP 2025 Administration: English I

Group	Grade							
	6	7	8	9	10	11	12	Total
All Students	<10	<10	≥2,050	≥40,160	≥3,200	≥880	≥420	≥46,730
Gender								
Female	<10	<10	≥1,130	≥19,840	≥1,110	≥300	≥140	≥22,540
Male	<10	<10	≥910	≥20,310	≥2,080	≥580	≥270	≥24,180
Ethnicity								
Hispanic/Latino	<10	<10	≥280	≥2,850	≥520	≥190	≥100	≥3,970
American Indian or Alaska Native	<10	<10	<10	≥270	≥30	<10	<10	≥320
Asian	<10	<10	140	≥600	≥30	≥10	<10	≥790
Black or African American	<10	<10	≥630	≥16,750	≥1,620	≥510	≥260	≥19,790
Native Hawaiian or Other Pacific	<10	<10	<10	≥30	<10	<10	<10	≥30
White	<10	<10	≥900	≥18,570	≥930	≥140	≥40	≥20,590
Two or More Races	<10	<10	≥80	≥1,060	≥40	≥10	<10	≥1,210
Education Classification								
Regular	<10	<10	≥1,550	≥34,390	≥2,410	≥640	≥330	≥39,340
Special	<10	<10	≥30	≥3,610	≥760	≥230	≥80	≥4,730
Gifted	<10	<10	≥460	≥2,150	≥20	≥10	<10	≥2,650
Economic Status*								
Economically Disadvantaged	<10	<10	≥1,070	≥25,870	≥2,620	≥680	≥160	≥30,430
Not Economically Disadvantaged	<10	<10	≥840	≥11,870	≥420	≥60	<10	≥13,200
English Learner Status								
Non-EL	<10	<10	≥2,040	≥39,120	≥2,750	≥690	≥310	≥44,920
EL	<10	<10	≥10	≥1,040	≥440	≥190	≥100	≥1,800
Migrant Status								
Nonmigrant	<10	<10	≥2,050	≥40,110	≥3,180	≥880	≥410	≥46,650
Migrant	<10	<10	<10	≥40	≥10	<10	<10	≥70
Section 504 Status								
Non-Section 504	<10	<10	≥1,940	≥36,540	≥2,660	≥760	≥360	≥42,280
Section 504	<10	<10	≥110	≥3,610	≥540	≥120	≥50	≥4,450
Homeless Status								
Not Homeless	<10	<10	≥2,050	≥39,640	≥3,110	≥850	≥410	≥46,080
Homeless	<10	<10	<10	≥510	≥80	≥20	<10	≥640
Military Affiliation								
Not Military Affiliated	<10	<10	≥2,000	≥39,620	≥3,170	≥880	≥410	≥46,100
Military Affiliated	<10	<10	≥50	≥530	≥20	<10	<10	≥620
Foster Care Status								
Not in Foster Care	<10	<10	≥2,050	≥40,030	≥3,170	≥880	≥410	≥46,570
Foster Care	<10	<10	≥120	≥20	<10	<10	≥160	<10

*Economic Status was not available for all students.

Table 7.6 Reportable Percentage of Students taking the Spring 2021 LEAP 2025 Administration: English I

Group	Grade							
	6	7	8	9	10	11	12	Total
All Students	0.00	0.01	4.40	85.94	6.85	1.90	0.90	100
Gender								
Female	0.00	0.02	5.05	88.01	4.94	1.35	0.63	100
Male	0.00	0.01	3.79	84.01	8.63	2.41	1.15	100
Ethnicity								
Hispanic/Latino	0.00	0.15	7.25	71.84	13.17	4.99	2.59	100
American Indian or Alaska Native	0.00	0.00	2.46	85.54	9.85	1.23	0.92	100
Asian	0.00	0.00	17.61	75.72	4.53	1.89	0.25	100
Black or African American	0.00	0.01	3.19	84.63	8.22	2.61	1.34	100
Native Hawaiian or Other Pacific	0.00	0.00	2.70	91.89	5.41	0.00	0.00	100
White	0.00	0.00	4.39	90.21	4.52	0.68	0.20	100
Two or More Races	0.00	0.00	6.85	87.62	4.04	1.07	0.41	100
Education Classification								
Regular	0.00	0.00	3.96	87.42	6.15	1.63	0.85	100
Special	0.00	0.00	0.72	76.38	16.10	4.99	1.82	100
Gifted	0.00	0.23	17.52	81.12	0.75	0.38	0.00	100
Economic Status*								
Economically Disadvantaged	0.00	0.00	3.55	85.03	8.64	2.24	0.54	100
Not Economically Disadvantaged	0.00	0.01	6.37	89.88	3.20	0.48	0.07	100
English Learner Status								
Non-EL	0.00	0.00	4.55	87.07	6.14	1.54	0.69	100
EL	0.00	0.33	0.78	57.68	24.51	10.65	6.05	100
Migrant Status								
Nonmigrant	0.00	0.00	4.41	85.97	6.83	1.89	0.90	100
Migrant	0.00	8.33	0.00	68.06	16.67	5.56	1.39	100
Section 504 Status								
Non-Section 504	0.00	0.00	4.60	86.44	6.29	1.80	0.86	100
Section 504	0.00	0.13	2.47	81.17	12.15	2.83	1.24	100
Homeless Status								
Not Homeless	0.00	0.00	4.46	86.02	6.76	1.86	0.89	100
Homeless	0.00	0.93	0.31	80.06	13.08	4.21	1.40	100
Military Affiliation								
Not Military Affiliated	0.00	0.00	4.34	85.94	6.89	1.92	0.91	100
Military Affiliated	0.00	0.95	8.59	85.69	4.13	0.48	0.16	100
Foster Care Status								
Not in Foster Care	0.00	0.02	4.40	85.96	6.83	1.89	0.90	100
Foster Care	0.00	0.00	3.73	79.50	13.66	2.48	0.62	100

*Economic Status was not available for all students.

Table 7.7 Count of Students taking the Spring 2021 LEAP 2025 Administration: English II

Group	Grade							
	6	7	8	9	10	11	12	Total
All Students	<10	<10	<10	≥2,180	≥35,440	≥2520	≥680	≥40,830
Gender								
Female	<10	<10	<10	≥1,140	≥17,930	≥870	≥230	≥20,180
Male	<10	<10	<10	≥1,030	≥17,510	≥1640	≥450	≥20,650
Ethnicity								
Hispanic/Latino	<10	<10	<10	≥230	≥2,270	≥310	≥150	≥2,970
American Indian or Alaska Native	<10	<10	<10	<10	≥240	≥20	<10	≥280
Asian	<10	<10	<10	≥70	≥550	≥30	<10	≥660
Black or African American	<10	<10	<10	≥850	≥14,560	≥1,310	≥400	≥17,130
Native Hawaiian or Other Pacific	<10	<10	<10	<10	≥30	<10	<10	≥30
White	<10	<10	<10	≥920	≥16,960	≥780	≥100	≥18,780
Two or More Races	<10	<10	<10	≥70	≥820	≥40	≥10	≥950
Education Classification								
Regular	<10	<10	<10	≥1,730	≥30,660	≥1,890	≥520	≥34,810
Special	<10	<10	<10	≥120	≥2,790	≥600	≥150	≥3,670
Gifted	<10	<10	<10	≥320	≥1,990	≥10	<10	≥2,340
Economic Status*								
Economically Disadvantaged	<10	<10	<10	≥1,400	≥21,940	≥2,010	≥370	≥25,730
Not Economically Disadvantaged	<10	<10	<10	≥660	≥11,850	≥330	≥30	≥12,880
English Learner Status								
Non-EL	<10	<10	<10	≥2,130	≥34,650	≥2,240	≥520	≥39,570
EL	<10	<10	<10	≥40	≥790	≥270	≥150	≥1,260
Migrant Status								
Nonmigrant	<10	<10	<10	≥2,170	≥35,410	≥2,510	≥680	≥40,790
Migrant	<10	<10	<10	<10	≥30	≥10	<10	≥40
Section 504 Status								
Non-Section 504	<10	<10	<10	≥2,010	≥32,340	≥2,120	≥570	≥37,060
Section 504	<10	<10	<10	≥160	≥3,100	≥390	≥100	≥3,770
Homeless Status								
Not Homeless	<10	<10	<10	≥2,140	≥35,030	≥2,470	≥660	≥40,310
Homeless	<10	<10	<10	≥30	≥410	≥40	≥10	≥510
Military Affiliation								
Not Military Affiliated	<10	<10	<10	≥2,150	≥34,970	≥2,510	≥680	≥40,330
Military Affiliated	<10	<10	<10	≥20	≥470	<10	<10	≥500
Foster Care Status								
Not in Foster Care	<10	<10	<10	≥2,170	≥35,340	≥2,500	≥680	≥40,700
Foster Care	<10	<10	<10	≥10	≥100	≥10	<10	≥120

*Economic Status was not available for all students.

Table 7.8 Reportable Percentage of Students taking the Spring 2021 LEAP 2025 Administration: English II

Group	Grade							
	6	7	8	9	10	11	12	Total
All Students	0.00	0.00	0.00	5.34	86.81	6.17	1.67	100
Gender								
Female	0.00	0.00	0.00	5.66	88.87	4.33	1.14	100
Male	0.00	0.00	0.00	5.02	84.80	7.98	2.20	100
Ethnicity								
Hispanic/Latino	0.00	0.00	0.00	7.93	76.44	10.52	5.11	100
American Indian or Alaska Native	0.00	0.00	0.00	3.14	83.97	10.10	2.79	100
Asian	0.00	0.00	0.15	11.69	82.46	4.65	1.05	100
Black or African American	0.00	0.00	0.01	4.98	84.99	7.66	2.36	100
Native Hawaiian or Other Pacific	0.00	0.00	0.00	6.06	90.91	3.03	0.00	100
White	0.00	0.00	0.00	4.95	90.34	4.18	0.54	100
Two or More Races	0.00	0.00	0.00	7.55	86.16	5.14	1.15	100
Education Classification								
Regular	0.00	0.00	0.01	4.99	88.06	5.44	1.51	100
Special	0.00	0.00	0.00	3.27	75.94	16.58	4.22	100
Gifted	0.00	0.00	0.00	13.79	85.28	0.77	0.17	100
Economic Status*								
Economically Disadvantaged	0.00	0.00	0.00	5.44	85.27	7.81	1.47	100
Not Economically Disadvantaged	0.00	0.00	0.00	5.15	92.00	2.59	0.26	100
English Learner Status								
Non-EL	0.00	0.00	0.01	5.40	87.57	5.68	1.34	100
EL	0.00	0.00	0.00	3.40	62.79	21.54	12.27	100
Migrant Status								
Nonmigrant	0.00	0.00	0.00	5.34	86.82	6.16	1.67	100
Migrant	0.00	0.00	0.00	2.27	72.73	22.73	2.27	100
Section 504 Status								
Non-Section 504	0.00	0.00	0.01	5.44	87.26	5.74	1.55	100
Section 504	0.00	0.00	0.00	4.37	82.36	10.41	2.86	100
Homeless Status								
Not Homeless	0.00	0.00	0.00	5.32	86.88	6.14	1.66	100
Homeless	0.00	0.00	0.00	7.16	80.85	8.90	3.09	100
Military Affiliation								
Not Military Affiliated	0.00	0.00	0.00	5.35	86.72	6.23	1.69	100
Military Affiliated	0.00	0.00	0.00	4.15	93.68	1.58	0.59	100
Foster Care Status								
Not in Foster Care	0.00	0.00	0.00	5.33	86.83	6.16	1.67	100
Foster Care	0.00	0.00	0.00	7.81	80.47	9.38	2.34	100

*Economic Status was not available for all students.

Table 7.9 Count of Students taking the Spring 2021 LEAP 2025 Administration: Algebra I

Group	Grade							
	6	7	8	9	10	11	12	Total
All Students	<10	≥210	≥7,020	≥34,890	≥5,180	≥920	≥250	≥48,500
Gender								
Female	<10	≥100	≥3,740	≥17,270	≥2,090	≥380	≥110	≥23,720
Male	<10	≥110	≥3,270	≥17,620	≥3,090	≥540	≥130	≥24,780
Ethnicity								
Hispanic/Latino	<10	≥10	≥550	≥2,670	≥480	≥120	≥30	≥3,880
American Indian or Alaska Native	<10	<10	≥20	≥230	≥50	<10	<10	≥320
Asian	<10	≥30	≥310	≥360	≥30	≥10	<10	≥760
Black or African American	<10	≥50	≥1,750	≥15,530	≥2,900	≥540	≥170	≥20,950
Native Hawaiian or Other Pacific	<10	<10	<10	≥30	<10	<10	<10	≥40
White	<10	≥100	≥4,170	≥15,050	≥1,600	≥220	≥40	≥21,190
Two or More Races	<10	<10	≥200	≥1,000	≥100	≥10	<10	≥1,330
Education Classification								
Regular	<10	≥120	≥5,410	≥30,310	≥4,010	≥700	≥180	≥40,760
Special	<10	<10	≥140	≥3,280	≥1,120	≥210	≥60	≥4,820
Gifted	<10	≥80	≥1,470	≥1,290	≥40	<10	<10	≥2,910
Economic Status*								
Economically Disadvantaged	<10	≥50	≥3,120	≥23,740	≥4,200	≥710	≥120	≥31,960
Not Economically Disadvantaged	<10	≥140	≥3,210	≥9,080	≥720	≥90	≥10	≥13,270
English Learner Status								
Non-EL	<10	≥210	≥6,950	≥33,860	≥4,820	≥820	≥220	≥46,910
EL	<10	<10	≥60	≥1,030	≥360	≥100	≥20	≥1,590
Migrant Status								
Nonmigrant	<10	≥210	≥7,020	≥34,850	≥5,170	≥920	≥250	≥48,430
Migrant	<10	<10	<10	≥40	≥10	<10	<10	≥60
Section 504 Status								
Non-Section 504	<10	≥200	≥6,710	≥31,480	≥4,430	≥780	≥220	≥43,850
Section 504	<10	<10	≥310	≥3,410	≥750	≥140	≥30	≥4,650
Homeless Status								
Not Homeless	<10	≥210	≥6,990	≥34,390	≥5,060	≥900	≥240	≥47,820
Homeless	<10	<10	≥30	≥490	≥120	≥20	<10	≥670
Military Affiliation								
Not Military Affiliated	<10	≥200	≥6,850	≥34,460	≥5,150	≥920	≥240	≥47,850
Military Affiliated	<10	<10	≥170	≥430	≥30	<10	<10	≥650
Foster Care Status								
Not in Foster Care	<10	≥210	≥7,010	≥34,780	≥5,140	≥910	≥250	≥48,320
Foster Care	<10	<10	≥10	≥110	≥40	<10	<10	≥170

*Economic Status was not available for all students.

Table 7.10 Reportable Percentage of Students taking the Spring 2021 LEAP 2025 Administration: Algebra I

Group	Grade							
	6	7	8	9	10	11	12	Total
All Students	0.01	0.44	14.48	71.95	10.69	1.91	0.52	100
Gender								
Female	0.01	0.42	15.80	72.83	8.84	1.61	0.48	100
Male	0.02	0.45	13.22	71.09	12.47	2.19	0.55	100
Ethnicity								
Hispanic/Latino	0.00	0.31	14.23	68.80	12.54	3.27	0.85	100
American Indian or Alaska Native	0.00	0.31	7.50	72.50	16.56	2.50	0.63	100
Asian	0.52	4.44	40.47	48.17	4.96	1.31	0.13	100
Black or African American	0.00	0.28	8.35	74.12	13.84	2.58	0.82	100
Native Hawaiian or Other Pacific	0.00	4.17	16.67	64.58	10.42	4.17	0.00	100
White	0.01	0.47	19.69	71.02	7.55	1.07	0.19	100
Two or More Races	0.00	0.30	15.55	75.28	7.74	0.90	0.23	100
Education Classification								
Regular	0.00	0.31	13.28	74.37	9.85	1.72	0.46	100
Special	0.00	0.04	2.90	68.08	23.25	4.48	1.24	100
Gifted	0.21	2.88	50.51	44.37	1.65	0.27	0.10	100
Economic Status*								
Economically Disadvantaged	0.00	0.18	9.77	74.27	13.16	2.23	0.39	100
Not Economically Disadvantaged	0.00	1.11	24.22	68.39	5.42	0.75	0.11	100
English Learner Status								
Non-EL	0.01	0.45	14.83	72.19	10.28	1.76	0.48	100
EL	0.00	0.00	4.33	64.76	22.82	6.33	1.76	100
Migrant Status								
Nonmigrant	0.01	0.44	14.50	71.95	10.67	1.91	0.52	100
Migrant	0.00	0.00	2.94	67.65	25.00	1.47	2.94	100
Section 504 Status								
Non-Section 504	0.01	0.47	15.31	71.80	10.11	1.79	0.51	100
Section 504	0.00	0.17	6.73	73.29	16.16	3.01	0.64	100
Homeless Status								
Not Homeless	0.01	0.44	14.62	71.92	10.58	1.89	0.52	100
Homeless	0.00	0.00	4.57	73.45	18.44	2.95	0.59	100
Military Affiliation								
Not Military Affiliated	0.01	0.43	14.32	72.02	10.77	1.93	0.52	100
Military Affiliated	0.15	0.61	26.65	66.62	4.90	0.61	0.46	100
Foster Care Status								
Not in Foster Care	0.01	0.44	14.51	71.97	10.65	1.90	0.52	100
Foster Care	0.00	0.00	7.30	65.17	22.47	4.49	0.56	100

*Economic Status was not available for all students.

Table 7.11 Count of Students taking the Spring 2021 LEAP 2025 Administration: Geometry

Group	Grade							
	6	7	8	9	10	11	12	Total
All Students	<10	≥20	≥240	≥6,140	≥23,360	≥4,160	≥310	≥34,260
Gender								
Female	<10	≥10	≥130	≥3,330	≥12,430	≥2,160	≥150	≥18,230
Male	<10	<10	≥110	≥2,800	≥10,930	≥1,990	≥160	≥16,020
Ethnicity								
Hispanic/Latino	<10	<10	≥10	≥350	≥1,550	≥370	≥30	≥2,330
American Indian or Alaska Native	<10	<10	<10	≥20	≥150	≥20	<10	≥200
Asian	<10	<10	≥30	≥270	≥320	≥30	<10	≥680
Black or African American	<10	<10	≥60	≥1,550	≥9,760	≥2,470	≥210	≥14,070
Native Hawaiian or Other Pacific	<10	<10	<10	<10	≥10	<10	<10	≥20
White	<10	<10	≥120	≥3,780	≥10,980	≥1,180	≥50	≥16,130
Two or More Races	<10	<10	<10	≥130	≥570	≥60	<10	≥780
Education Classification								
Regular	<10	<10	≥140	≥4,930	≥20,970	≥3,610	≥270	≥29,950
Special	<10	<10	<10	≥80	≥1,380	≥410	≥30	≥1,920
Gifted	<10	≥10	≥100	≥1,120	≥1,000	≥130	<10	≥2,380
Economic Status*								
Economically Disadvantaged	<10	<10	≥50	≥2,570	≥14,550	≥3,170	≥210	≥20,580
Not Economically Disadvantaged	<10	≥20	≥180	≥3,280	≥7,670	≥820	≥20	≥12,010
English Learner Status								
Non-EL	<10	≥20	≥240	≥6,110	≥22,890	≥3,950	≥280	≥33,510
EL	<10	<10	<10	≥30	≥470	≥210	≥30	≥750
Migrant Status								
Nonmigrant	<10	≥20	≥240	≥6,130	≥23,350	≥4,160	≥310	≥34,230
Migrant	<10	<10	<10	<10	≥10	<10	<10	≥20
Section 504 Status								
Non-Section 504	<10	≥20	≥230	≥5,880	≥21,460	≥3,770	≥290	≥31,670
Section 504	<10	<10	≥10	≥250	≥1,900	≥380	≥20	≥2,590
Homeless Status								
Not Homeless	<10	≥20	≥240	≥6,100	≥23,100	≥4,080	≥300	≥33,860
Homeless	<10	<10	<10	≥30	≥260	≥80	≥10	≥390
Military Affiliation								
Not Military Affiliated	<10	≥20	≥240	≥6,000	≥23,040	≥4,110	≥310	≥33,740
Military Affiliated	<10	<10	<10	≥130	≥320	≥50	<10	≥510
Foster Care Status								
Not in Foster Care	<10	≥20	≥240	≥6,130	≥23,300	≥4,150	≥310	≥34,180
Foster Care	<10	<10	<10	<10	≥50	≥10	<10	≥80

*Economic Status was not available for all students.

Table 7.12 Percentage of Students taking the Spring 2021 Administration: Geometry

Group	Grade							
	6	7	8	9	10	11	12	Total
All Students	0.00	0.06	0.73	17.92	68.21	12.15	0.93	100
Gender								
Female	0.00	0.07	0.72	18.30	68.19	11.88	0.83	100
Male	0.01	0.06	0.73	17.49	68.22	12.46	1.03	100
Ethnicity								
Hispanic/Latino	0.00	0.00	0.51	15.33	66.42	16.23	1.50	100
American Indian or Alaska Native	0.00	0.00	0.49	12.81	75.86	10.84	0.00	100
Asian	0.15	1.16	5.09	40.26	47.38	5.67	0.29	100
Black or African American	0.00	0.01	0.49	11.05	69.32	17.59	1.53	100
Native Hawaiian or Other Pacific	0.00	3.45	6.90	20.69	62.07	6.90	0.00	100
White	0.00	0.05	0.76	23.45	68.03	7.35	0.37	100
Two or More Races	0.00	0.25	1.02	17.01	73.48	7.61	0.63	100
Education Classification								
Regular	0.00	0.01	0.48	16.47	70.04	12.07	0.93	100
Special	0.00	0.00	0.16	4.36	72.00	21.51	1.97	100
Gifted	0.04	0.80	4.28	47.09	42.10	5.70	0.00	100
Economic Status*								
Economically Disadvantaged	0.00	0.00	0.29	12.51	70.70	15.44	1.05	100
Not Economically Disadvantaged	0.00	0.17	1.51	27.38	63.87	6.84	0.22	100
English Learner Status								
Non-EL	0.00	0.06	0.74	18.24	68.31	11.79	0.85	100
EL	0.00	0.00	0.00	4.00	63.33	28.27	4.40	100
Migrant Status								
Nonmigrant	0.00	0.06	0.73	17.93	68.20	12.16	0.93	100
Migrant	0.00	0.00	0.00	9.09	81.82	9.09	0.00	100
Section 504 Status								
Non-Section 504	0.00	0.06	0.74	18.58	67.77	11.92	0.93	100
Section 504	0.00	0.04	0.62	9.92	73.51	15.02	0.89	100
Homeless Status								
Not Homeless	0.00	0.06	0.74	18.03	68.22	12.06	0.90	100
Homeless	0.00	0.00	0.00	9.05	67.34	20.35	3.27	100
Military Affiliation								
Not Military Affiliated	0.00	0.06	0.71	17.81	68.29	12.18	0.94	100
Military Affiliated	0.00	0.00	1.54	25.68	62.55	10.23	0.00	100
Foster Care Status								
Not in Foster Care	0.00	0.06	0.73	17.95	68.19	12.14	0.93	100
Foster Care	0.00	0.00	0.00	8.75	73.75	17.50	0.00	100

*Economic Status was not available for all students.

Tables 7.13 through 7.16 summarize the mean scale scores, standard deviations, and the percentage of students in each achievement level for the 2020–2021 administration of the LEAP 2025 high school ELA and mathematics assessments. All three administrations are presented

Table 7.13 Comparison of Percentage of Students in Each Achievement Level: English I

	Year	Administration	Form		Scale Score		Percentage in Achievement Level				
				N	Mean	SD	1	2	3	4	5
All	2020	Fall	A	≥7,070	732.15	37.04	21.6	22.8	23.1	25.5	7.1
	2021	Spring	E	≥46,730	739.38	36.37	14.3	20.2	25.9	32.3	7.3
	2021	Summer	B	≥2,580	695.97	24.07	58.0	30.6	9.1	2.2	0.2
First-Time Testers	2020	Fall	A	≥5,720	740.73	34.86	12.3	20.8	27.1	31.2	8.7
	2021	Spring	E	≥45,150	740.94	35.70	12.7	19.9	26.6	33.3	7.5
	2021	Summer	B	≥360	714.17	30.76	34.0	29.3	23.6	12.1	1.1
Retesters	2020	Fall	A	≥1,200	693.36	18.01	65.7	29.7	4.1	0.5	0.0
	2021	Spring	E	≥1,230	688.17	19.89	69.0	27.9	3.2	0.0	0.0
	2021	Summer	B	≥2,130	692.40	21.08	62.8	30.5	6.1	0.5	0.0
Previously Passed	2020	Fall	A	≥150	716.58	25.61	21.4	43.5	24.0	9.7	1.3
	2021	Spring	E	≥340	718.43	28.82	23.5	35.8	27.0	13.2	0.6
	2021	Summer	B	≥70	708.79	22.82	36.4	40.3	22.1	1.3	0.0

Levels: 1 = *Unsatisfactory*, 2 = *Approaching Basic*, 3 = *Basic*, 4 = *Mastery*, 5 = *Advanced*

Table 7.14 Comparison of Percentage of Students in Each Achievement Level: English II

	Year	Administration	Form		Scale Score		Percentage in Achievement Level				
				N	Mean	SD	1	2	3	4	5
All	2020	Fall	A	≥8,880	736.70	42.27	20.5	19.6	22.1	28.5	9.2
	2021	Spring	E	≥40,830	743.93	46.18	18.1	16.3	21.1	29.0	15.5
	2021	Summer	B	≥2,640	691.45	28.68	62.6	26.2	8.3	2.5	0.4
First-Time Testers	2020	Fall	A	≥8,280	740.27	40.99	16.7	19.7	23.3	30.5	9.9
	2021	Spring	E	≥39,510	745.81	45.46	16.3	16.2	21.6	29.9	16.0
	2021	Summer	B	≥270	711.69	42.11	41.9	21.7	15.9	17.3	3.2
Retesters	2020	Fall	A	≥550	684.17	22.61	78.6	17.4	3.3	0.7	0.0
	2021	Spring	E	≥1,160	682.45	23.40	77.7	17.4	3.9	1.0	0.0
	2021	Summer	B	≥2,350	688.93	25.52	65.1	26.8	7.4	0.7	0.0
Previously Passed	2020	Fall	A	≥40	722.75	28.35	22.9	33.3	27.1	14.6	2.1
	2021	Spring	E	≥150	724.58	34.93	24.7	24.0	28.6	20.8	1.9
	2021	Summer	B	≥10	714.79	35.30	42.9	21.4	21.4	14.3	0.0

Levels: 1 = *Unsatisfactory*, 2 = *Approaching Basic*, 3 = *Basic*, 4 = *Mastery*, 5 = *Advanced*

Table 7.15 Comparison of Percentage of Students in Each Achievement Level: Algebra I

	Year	Administration	Form		Scale Score		Percentage in Achievement Level				
				N	Mean	SD	1	2	3	4	5
All	2020	Fall	D	≥5,100	728.25	30.97	17.7	30.1	28.2	22.8	1.2
	2021	Spring	E	≥48,500	733.34	34.40	17.2	27.2	26.4	26.2	3.0
	2021	Summer	D	≥3,540	708.23	22.52	34.9	45.4	16.4	2.7	0.5
First-Time Testers	2020	Fall	D	≥4,290	732.22	31.04	14.1	27.2	30.6	26.6	1.4
	2021	Spring	E	≥47,320	734.02	34.34	16.6	26.9	26.7	26.7	3.1
	2021	Summer	D	≥420	725.79	36.25	23.5	33.8	21.9	16.6	4.2
Retesters	2020	Fall	D	≥660	704.55	18.06	39.7	46.5	12.8	0.9	0.0
	2021	Spring	E	≥730	698.43	19.49	54.4	36.7	8.8	0.1	0.0
	2021	Summer	D	≥3,060	705.60	18.46	36.8	47.1	15.3	0.8	0.0
Previously Passed	2020	Fall	D	≥140	718.72	24.77	21.4	40.0	28.3	10.3	0.0
	2021	Spring	E	≥450	718.29	25.29	23.9	38.1	26.4	11.3	0.2
	2021	Summer	D	≥40	718.86	24.66	20.4	40.8	30.6	6.1	2.0

Levels: 1 = *Unsatisfactory*, 2 = *Approaching Basic*, 3 = *Basic*, 4 = *Mastery*, 5 = *Advanced*

Table 7.16 Comparison of Percentage of Students in Each Achievement Level: Geometry

	Year	Administration	Form		Scale Score		Percentage in Achievement Level				
				N	Mean	SD	1	2	3	4	5
All	2020	Fall	D	≥5,680	733.37	26.70	6.9	34.3	31.4	23.2	4.3
	2021	Spring	E	≥34,260	734.21	26.60	6.2	32.6	32.6	24.6	4.0
	2021	Summer	D	≥1,050	711.27	17.30	18.3	65.5	14.0	1.9	0.3
First-Time Testers	2020	Fall	D	≥5,600	733.61	26.69	6.7	34.0	31.6	23.4	4.4
	2021	Spring	E	≥33,920	734.41	26.59	6.1	32.3	32.7	24.8	4.0
	2021	Summer	D	≥200	716.35	22.74	18.5	51.0	23.0	6.0	1.5
Retesters	2020	Fall	D	≥40	707.78	16.39	22.0	63.4	14.6	0.0	0.0
	2021	Spring	E	≥190	707.30	15.09	23.7	64.9	11.3	0.0	0.0
	2021	Summer	D	≥840	709.93	15.48	18.5	69.0	11.6	0.8	0.0
Previously Passed	2020	Fall	D	≥30	725.64	23.15	10.3	48.7	23.1	17.9	0.0
	2021	Spring	E	≥130	721.59	19.10	8.7	48.6	37.0	5.1	0.7
	2021	Summer	D	≥10	718.25	17.22	6.3	62.5	25.0	6.3	0.0

Levels: 1 = *Unsatisfactory*, 2 = *Approaching Basic*, 3 = *Basic*, 4 = *Mastery*, 5 = *Advanced*

7.2 Reports

Score reports are the primary means of communicating test scores to appropriate school system personnel (e.g., testing coordinators or superintendents), teachers, and parents. Standard 6.10 of the *Standards* states:

> When test score information is released, those responsible for testing programs should provide interpretations appropriate to the audience. The interpretations should describe in simple language what the test covers, what scores represent, the precision/reliability of the scores, and how scores are intended to be used (119).

Standard 5.1 is related to Standard 6.10. It states:

> Test users should be provided with clear explanations of the characteristics, meaning, and intended interpretation of scale scores, as well as their limitations (102).

Interpretations of test scores from each administration are disseminated in two ways: the individual score report and the *LEAP Interpretive Guide*.

In addition to providing interpretations of test results, the LDOE and DRC must ensure that those interpretations are understandable for the target audience. Standard 7.0 states:

> Information relating to tests should be clearly documented so that those who use tests can make informed decisions regarding which test to use for a specific purpose, how to administer the chosen test, and how to interpret test scores (125).

The LDOE and DRC strive to create documents that will be accessible to parents, teachers, and all other stakeholders.

The Individual Student-Level Report (ISR) is the primary means for sharing student test results with parents. As such, it is a stand-alone document from which parents can glean information that is relevant to

understanding their children's test scores. For more information about the test, parents are provided Parent Guide to the LEAP 2025 Student Reports. In the 2020–2021 administration year, student reports for each school were posted by subject, then downloaded and printed from eDIRECT by the school systems and schools. eDIRECT is DRC's secure online system that provides schools and districts access to student tests and reports.

7.3 Description of Each Type of Report

In this section, descriptions of the School Roster Report and the ISR are provided.

In compliance with AERA, APA, & NCME (2014) Standard 12.18, the LEAP 2025 score reports provide clear information about the results of individual students and of specific groups of students. Standard 12.18 states:

> In educational settings, score reports should be accompanied by a clear presentation of information on how to interpret the scores, including the degree of measurement error associated with each score or classification level, and by supplementary information related to group summary scores. In addition, dates of test administration and relevant norming studies should be included in score reports (200).

School Roster Report

A School Roster Report, which provides summary information about student performance on the LEAP 2025 high school ELA and mathematics assessments, is available to school systems and schools through eDIRECT. Total test scores and achievement level indicators are shown for the test of interest. Category and subcategory performance ratings are also reported for students. At the school level, the percentage of students at each achievement level and rating by category and subcategory are summarized. More details can be found in the LEAP Interpretive Guide.

Individual Student-Level Report

The ISR is another type of report available through the eDIRECT system. ISRs may be downloaded and printed by schools to be sent home to parents. At the top of the page, overall student performance is reported by scale score and achievement level. In the middle of the page, category and subcategory performance indicators are reported. When a student does not receive a scale score, their achievement level will be left blank. ISRs for students whose scores were invalidated will display a blank scale score for a given course.

A data file referred to as Louisiana Department of Education Student File (LDESTD) was provided to LDOE by DRC. It contains one record for every student tested; each record contains demographic information, responses for multiple-choice (MC) items, scores for items that are not MC items, raw scores, content and process standard raw scores, scale scores, and performance-level data for each content area.

The LEAP Interpretive Guide was written to help Louisiana school system and school administrators, teachers, parents, and the general public understand the LEAP 2025 ELA and mathematics tests. The *LEAP Interpretive Guide* was developed collaboratively by DRC and LDOE staff. LDOE staff had opportunities to review the guide, provide feedback, and give final approval.

The LEAP Interpretive Guide has three sections. The first section presents an introduction and an overview of key terms and test-related concepts. The second section discusses assessment terms and types of scores that are presented on the ISRs. Sample ISRs are included in the guide. The third section discusses information that is presented on the School Roster Report and an example of the report.

7.4 Summary

In summary, the overall purpose of reporting test results is to communicate information on student performance to stakeholders. These results are presented in the context of score reports that aid the user in

understanding the meaning of the test scores. The reports and ancillary information developed by DRC address multiple best practices of the testing industry but are particularly related to the following standards:

> *Standard 5.1* Test users should be provided with clear explanations of the characteristics, meaning, and intended interpretation of scale scores, as well as their limitations (102).
>
> *Standard 6.10* When test score information is released, those responsible for testing programs should provide interpretations appropriate to the audience. The interpretations should describe in simple language what the test covers, what scores represent, the precision/reliability of the scores, and how scores are intended to be used (119).
>
> *Standard 7.0* Information relating to tests should be clearly documented so that those who use tests can make informed decisions regarding which test to use for a specific purpose, how to administer the chosen test, and how to interpret test scores (125).
>
> *Standard 12.18* In educational settings, score reports should be accompanied by a clear presentation of information on how to interpret the scores, including the degree of measurement error associated with each score or classification level, and by supplementary information related to group summary scores. In addition, dates of test administration and relevant norming studies should be included in score reports (200).

Chapter 8: Performance-Level Setting

This chapter briefly describes the LEAP 2025 high school performance-level setting and presents the cut scores and achievement-level descriptors derived from the performance-level setting. Since the LDOE used PARCC cut scores for the LEAP 2025 high school tests, a brief overview of the PARCC performance-level setting procedures is included in this chapter. A more detailed discussion and the results of the PARCC performance-level setting may be found in the *Performance Level Setting Technical Report* (Pearson, 2015).

The AERA, APA, & NCME (2014) *Standards* addressed by the *Performance Level Setting Technical Report* (Pearson, 2015) are 5.21 and 5.22.

Starting in the 2017–2018 administrations, the LEAP 2025 High School assessments measured different content and constructs than did previous tests administered to Louisiana students. The new tests were built using the PARCC item bank and were fully aligned to the Louisiana Student Standards. The new tests were reported on new scales, and students were classified by achievement levels based on their knowledge and ability to perform different tasks in relation to the new test content and standards.

In terms of the validity of the LEAP 2025 test scores, it is essential to understand that descriptors and cut scores are established in a collaborative and participatory process. The descriptors clearly establish, in plain language, the proper frame of reference for understanding how to interpret test scores, particularly cut scores.

8.1 PARCC Performance-Level Setting Process for English Language Arts and Mathematics

According to the *Performance Level Setting Technical Report* (Pearson, 2015), PARCC used the evidence-based standard setting (EBSS) method (Beimers, Way, McClarty, & Miles, 2012) for the PARCC performance-level setting (PLS) process. The EBSS method is used to combine various considerations into the process for setting performance levels, including policy considerations, content standards, research, and educator judgment about what students should know and be able to demonstrate to support PARCC's policy goals related to college- and career-readiness expectations. Additional details about the EBSS method can be found in the *Performance Level Setting Technical Report* (Pearson, 2015).

8.2 Cut Scores

This section presents the cut scores for each grade and content area of the LEAP 2025 High School assessments. Table 8.1 summarizes the cut scores.

Table 8.1 LEAP 2025 High School Assessment Cut Scores

Content	Approaching Basic	Basic	Mastery	Advanced
English I	700	725	750	791
English II	700	725	750	794
Algebra I	700	725	750	805
Geometry	700	725	750	783

8.3 Category Cut Scores

As stated in Chapter 6, section "Category- and Subcategory-Level Subscores," student performance on ELA and mathematics reporting categories and subcategories was classified into one of three performance ratings: *Strong*, *Moderate*, and *Weak*. Detailed rules for calculating performance ratings for ELA and mathematics categories and subcategories can be found in that section.

8.4 Achievement-Level Definitions

The cut scores divide the continuum of student achievement into the following five achievement levels used by LDOE for reporting purposes:

- *Advanced*: Students performing at this level have **exceeded** college- and career-readiness expectations and are well prepared for the next level of study in this content area.
- *Mastery*: Students performing at this level have **met** college- and career-readiness expectations and are prepared for the next level of study in this content area.
- *Basic*: Students performing at this level have **nearly met** college- and career-readiness expectations and may need additional support to be fully prepared for the next level of study in this content area.
- *Approaching Basic*: Students performing at this level have **partially met** college- and career-readiness expectations and will need much support to be prepared for the next level of study in this content area.
- *Unsatisfactory*: Students performing at this level have **not yet met** the college- and career-readiness expectations and will need extensive support to be prepared for the next level of study in this content area.

Table 8.2 summarizes the LEAP 2025 High School scale-score ranges for each level of achievement.

Table 8.2 Achievement-Level Scale-Score Ranges

Content	Unsatisfactory	Approaching Basic	Basic	Mastery	Advanced
English I	650–699	700–724	725–749	750–790	791–850
English II	650–699	700–724	725–749	750–793	794–850
Algebra I	650–699	700–724	725–749	750–804	805–850
Geometry	650–699	700–724	725–749	750–782	783–850

8.5 Summary

This chapter presented a brief overview of PARCC's performance-level setting process, which set the cut scores used by LDOE for reporting student performance on the LEAP 2025 High School tests. These procedures are addressed in more detail in relevant technical reports.

The performance-level setting process undertaken by PARCC addresses the following standards:

> *Standard 5.21* When proposed score interpretations involve one or more cut scores, the rationale and procedures used for establishing cut scores should be documented clearly (107).
>
> *Standard 5.22* When cut scores defining pass-fail or proficiency levels are based on direct judgments about the adequacy of item or test performances, the judgmental process should be designed so that the participants providing the judgments can bring their knowledge and experience to bear in a reasonable way (108).

Chapter 9: Evidence of Construct-Related Validity

Evidence for construct-related validity—the meaning of test scores and the inferences they support—is the central concept underlying the LEAP 2025 validation process. Validity evidence, from the design of the test to item development and scoring, is created throughout the entire assessment process. Therefore, evidence of validity is described throughout the LEAP 2025 technical report. Table 9.1 summarizes the sources of validity evidence and indicates where the evidence can be found in the technical report.

Table 9.1 Evidence of Validity and the Corresponding Technical Report Chapter

Source of Validity	Related Information	Related Chapter/Source
Evidence Based on Test Content	Item Development Process	Chapter 3 2020-2021 LEAP 2025 High School ELA and Mathematics Assessment Frameworks
	Test Blueprint and Item Alignment to Curriculum and Standards	Chapter 3 2020-2021 LEAP 2025 High School ELA and Mathematics Assessment Frameworks
	Item Bias, Sensitivity, and Content Appropriateness	Chapter 3
	Accommodations	Chapters 3 and 4
Evidence Based on Response Processes	Data Review	2020-2021 LEAP 2025 High School ELA and Mathematics Assessment Frameworks
	Classical Item analysis	Chapter 6
Evidence Based on Internal Structure	Differential Item Functioning	Chapter 10
	Reliability and Standard Errors of Measurement	Chapter 9
Evidence Based on Relations to Other Variables	Divergent Validity	Chapter 9
Evidence Based on the Consequences of Testing	Scale Score and Performance Level Information	Chapter 7
	Test Interpretive Guide	Chapter 4

In this chapter, DRC presents evidence of construct-related validity through studies of test reliability, convergent validity, and divergent validity. All analyses in this chapter are based on the initial testers in the census data. Since the summer administration is made up of primarily a re-test population, summer results are not reported.

Chapter 9 of this report demonstrates adherence to the American Educational Research Association, American Psychological Association, & National Council on Measurement in Education (AERA, APA, & NCME, 2014) Standards 1.13, 1.21, 2.0, 2.3, 2.13, 2.14, 2.16, and 2.19. Each standard is discussed in the pertinent section of this chapter.

9.1 Construct-Irrelevant Variance and Construct Underrepresentation

Minimization of construct-irrelevant variance and construct underrepresentation is addressed in the following steps of the test development process: (1) specification, (2) item writing, (3) review, (4) field-testing, (5) test construction, and (6) item calibration (see Chapter 3 for more information on steps 1–5 and Chapter 6 for more information on step 6).

Construct-irrelevant variance refers to error variance that is caused by factors unrelated to the constructs measured by the test. For example, when tests are not administered under standardized conditions (e.g., one administration may be timed, but another administration is untimed), differences in student performance related to different administration conditions may result. Careful specification of content and review of the items representing that content are first steps in minimizing construct-irrelevant variance. Then, empirical evidence, especially item-level data, is used to infer construct irrelevance.

Construct underrepresentation occurs when the content of the assessment does not reflect the full range of content that the assessment is expected to cover. Specification and review, a process through which test blueprints are developed and reviewed, are primary steps in the development process designed to ensure that content is appropriately represented.

9.2 Reliability

Reliability refers to the consistency of students' test scores on parallel forms of a test. A reliable test is one that produces scores that are expected to be relatively stable if the test is administered repeatedly under similar conditions. Often, however, it is impractical to administer multiple forms of the test, and reliability is estimated on a single administration of the test. This type of reliability, known as internal consistency, provides an estimate of how consistently examinees perform across items within a test during a single test administration (Crocker & Algina, 1986). Reliability is a necessary, but not sufficient, condition of validity.

The 2014 *Standards* indicates the following:

> The term *reliability* has been used in two ways in the measurement literature. First, the term has been used to refer to the reliability coefficients of classical test theory, defined as the correlation between scores on two equivalent forms of the test, presuming that taking one form has no effect on performance on the second form. Second, the term has been used in a more general sense, to refer to the consistency of scores across replications of a testing procedure, regardless of how this consistency is estimated or reported (e.g., in terms of standard errors, reliability coefficients per se, generalizability coefficients, error/tolerance ratios, item response theory (IRT) information functions, or various indices of classification consistency) (33).

In accordance with the *Standards* in developing and maintaining tests of the highest quality, DRC has calculated the reliability of each LEAP 2025 test in a variety of ways: reliability of raw scores, overall

standard error of measurement (SEM), IRT-based conditional SEM, and decision consistency of achievement-level classifications.

There are several specific standards that this chapter addresses. These include Standards 2.0, 2.3, 2.13, and 2.19, each of which is articulated below.

> *Standard 2.0* Appropriate evidence of reliability/precision should be provided for the interpretation for each intended score use (42).
>
> *Standard 2.3* For each total score, subscore, or combination of scores that is to be interpreted, estimates of relevant indices of reliability/precision should be reported (43).

The total score reliabilities are discussed in Section 9.3 of this chapter. The category and subcategory reliabilities and SEMs are presented in Sections 9.11 and 9.4 and 9.11. The SEM of the total score is discussed in Section 9.4.

> *Standard 2.13* The standard error of measurement, both overall and conditional (if reported), should be provided in units of each reported score (45).

The SEM based on raw scores is discussed in Section 9.3 and is reported in raw score units. The conditional SEM is discussed in Section 9.5 and is presented in scale score units.

> *Standard 2.19* Each method of quantifying the reliability/precision of scores should be described clearly and expressed in terms of statistics appropriate to the method. The sampling procedures used to select test takers for reliability/precision analyses and the descriptive statistics on these samples, subject to privacy obligations where applicable, should be reported (47).

Section 9.3 discusses different ways of measuring test reliability, including reliability of raw scores and test-form SEM, IRT-based conditional SEM, and decision consistency of achievement-level classifications. These statistics were computed based on initial testers. Since the summer forms are primarily administered to students retesting, statistics for the summer form will not be reported. The summer form had been previously administered to a spring or fall population; therefore, the form's reliability information can be found in earlier technical reports.

9.3 Test Reliability

The reliability of raw scores by test form was evaluated using Cronbach's (1951) coefficient alpha, which is a lower-bound estimate of test reliability. The reliability coefficient is a ratio of the variance of true test scores to the variance of the total observed scores, with the values ranging from 0 to 1. The closer the value of the reliability coefficient is to 1, the more consistent the scores, where 1 refers to a perfectly consistent test. In general, reliability coefficients that are equal to or greater than 0.8 are considered acceptable for tests of moderate lengths.

Cronbach's coefficient alpha was computed using the formula

$$\alpha = \frac{n}{n-1}\left[1 - \frac{\sum_{i=1}^{n} \sigma_i^2}{\sigma_X^2}\right], \quad (9.1)$$

where n is the number of items on the test, σ_i^2 is the variance of item i, and σ_x^2 is the variance of the total test score.

Total test reliability measures, such as Cronbach's coefficient alpha and SEM, consider the consistency (i.e., reliability) of performance over all test questions in a given form, the results of which imply how well the questions measure the content domain and could continue to do so over repeated administrations. The number of items in the test influences these statistics; for example, a longer test can be expected to be more reliable than a shorter test.

The reliability coefficients for the fall and spring LEAP 2025 HS assessments are reported in Table 9.2. English I and English II have one writing component (RI or RL) that has the same score as another component (WE). The item score for the RI/RL writing component was excluded from the reliability computation. The reliability statistics ranged from 0.86 to 0.92 and from 0.90 to 0.91 for the fall and spring administrations, respectively. The two administrations had very similar reliability statistics. These results indicate acceptable reliability coefficients for the LEAP 2025 high school tests.

Table 9.2 Reliability

Administration	Course	Form	Number of Items	Number of Score Points	SEM	Cronbach's Alpha	N-Count
Fall 2020	English I	A	33	90	5.55	0.90	≥5,740
	English II	A	33	90	5.17	0.91	≥8,330
	Algebra I	D	39	68	3.58	0.87	≥4,290
	Geometry	D	39	68	3.52	0.91	≥5,610
Spring 2021	English I	E	33	90	5.39	0.91	≥45,840
	English II	E	33	90	5.50	0.91	≥39,950
	Algebra I	E	39	68	3.56	0.90	≥47,920
	Geometry	E	39	68	3.53	0.91	≥34,280

The reliability statistics by subgroup are reported and discussed in Chapter 10.

9.4 Standard Error of Measurement

The reliability of reported test scores can be characterized by the standard errors associated with the scores. The SEM may be used to determine the range within which a student's true score is likely to fall. An observed score should be regarded not as a student's true score but as an estimate of a student's true score. It is expected that the score a student obtains from a single test administration would fall within one SEM of the student's true score 68% of the time and within approximately two SEMs of the

true score 95% of the time. The SEM is an index of the random variability in test scores and is defined as follows:

$$\text{SEM} = SD\sqrt{1 - R_{xx'}}\ , \tag{9.2}$$

where SD represents standard deviation of the raw score distribution, and $R_{xx'}$ is estimated by $\hat{\alpha}$ as expressed in Equation 9.1.

The SEM at the test-form level was computed in raw score metric and is also presented in Table 9.2. With English I and English II, the raw score used to calculate the SD included the RI/RL component and weighting of WE.

9.5 Conditional Standard Error of Measurement

In contrast to SEM, conditional standard error of measurement (CSEM) expresses the degree of measurement error in scale score units and is conditioned on the ability of the student. DRC reports the CSEM in accordance with Standard 2.14, which states:

> When possible and appropriate, conditional standard errors of measurement should be reported at several score levels unless there is evidence that the standard error is constant across score levels. Where cut scores are specified for selection or classification, the standard errors of measurement should be reported in the vicinity of each cut score (46).

In further compliance with Standard 2.14, the CSEM of each cut score is reported in Table 9.3.

The CSEMs are defined as the reciprocal of the square root of the test information function and can be estimated across all points of the ability continuum (Hambleton & Swaminathan, 1985). The CSEM is defined in the following equation:

$$\text{CSEM}(\theta_i) = \frac{1}{\sqrt{I(\theta_i)}}\ , \tag{9.3}$$

where $I(\vartheta_i)$ is the test information function, as a sum of item information function 2, obtained as

$$I(\theta_i) = \sum_j \frac{p'_{ij}(\theta_i)^2}{p_{ij}(\theta_i)\, q_{ij}(\theta_i)}\ , \tag{9.4}$$

where $p'_{ij}(\theta_i)$ is the derivative of $p_{ij}(\theta_i)$ and $q_{ij}(\theta_i) = 1 - p_{ij}(\theta_i)$.

Note that the CSEMs vary in magnitude across the entire range of student ability estimates (i.e., scale scores) and are smaller in the middle of the score distribution and larger at the tails. This pattern is expected when IRT methods are used. Since LEAP 2025 was first administered, every effort has been made to make the TCC and CSEM values at the cut scores between the PARCC assessments and the LEAP

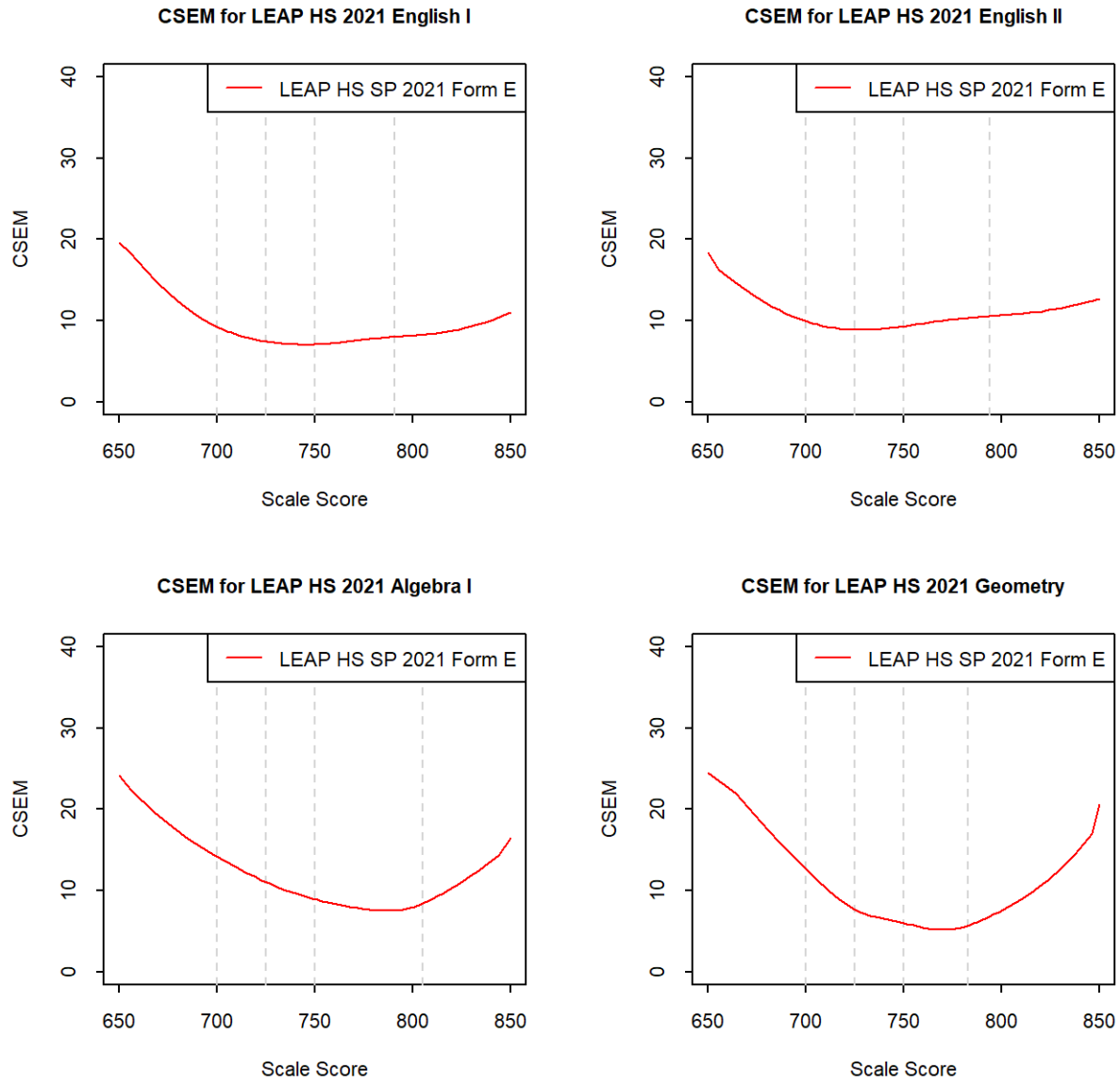
2025 assessments similar. Both TCC and CSEM values have been similar across the LEAP 2025 alternate forms given the same content because similar or the same statistical properties are important for alternate forms. To provide context regarding the magnitude of the CSEMs, it is important to also refer to sections 9.3 Test Reliability and 9.6 Classification Accuracy and Consistency where evidence is provided of high measures of form reliability and levels of accurate student classification at the cutpoints to support the use of the LEAP 2025 assessments. The CSEMs at the four cut scores that define the performance levels are presented in Table 9.3.

Table 9.3 Conditional Standard Errors of Measurement at the *Approaching Basic*, *Basic*, *Mastery*, and *Advanced* Cut Scores

			Approaching Basic		*Basic*		*Mastery*		*Advanced*	
Administration	**Course**	**Form**	**Cut Score**	**CSEM**	**Cut Score**	**CSEM**	**Cut Score**	**CSEM**	**Cut Score**	**CSEM**
Fall 2020	English I	A	700	10	725	8	750	8	791	9
	English II	A	700	11	725	10	750	10	794	12
	Algebra I	D	700	14	725	11	750	9	805	9
	Geometry	D	700	13	725	8	750	6	783	6
Spring 2021	English I	E	700	9	725	7	750	7	791	8
	English II	E	700	10	725	9	750	9	794	11
	Algebra I	E	700	14	725	11	750	9	805	8
	Geometry	E	700	13	725	8	750	6	783	6

Figure 9.1 displays the CSEM curves for each subject area. With fixed-form assessments, the estimates of measurement error tend to be higher at the low and high ends of the scale-score range, where few items measure the ability levels. Generally, there are few students with extreme scores, and these score levels cannot be estimated as accurately as levels toward the middle of the ability range. The middle of the ability range, where cut scores are located, shows lower measurement error than the low and high ends of the ability ranges. Figure 9.1 demonstrates that the tests are designed so that measurement error is minimized in the middle of the scale range, where most students are located.

Figure 9.1 CSEM Curves for LEAP High School 2021



9.6 Classification Accuracy and Consistency

Classification Accuracy

Classification accuracy is defined as the extent to which the actual classifications of test takers into various achievement levels match classifications made based on their true scores (Livingston & Lewis, 1995). Classification accuracy refers to the agreement between the observed score and the true score, whereas classification consistency refers to the agreement between two observed scores.

Classification Consistency

Classification consistency is defined as the extent to which the classifications of students in a particular achievement level match based on two independent administrations of the same test form or one administration of two parallel test forms. It is often logistically infeasible, as well as expensive, to obtain data from repeated administrations of a test, be it re-administration of the same test or administration of a parallel form. Therefore, a common practice is to estimate classification consistency from one administration of a test.

The Livingston-Lewis (1995) methodology was used to calculate classification accuracy statistics based on the spring 2019 LEAP 2025 results. The Livingston-Lewis procedure utilizes a beta-binomial model that requires two steps: (1) fitting proportion-correct true scores to a four-parameter beta distribution and (2) using the binomial distribution to estimate classification accuracy and consistency. All calculations for classification accuracy and consistency are based on census data.

Classification consistency and classification accuracy conditioned on achievement level (sees Table 9.4 and 9.5) and on cut score (see Tables 9.6 and 9.7) are presented for the 2019 LEAP 2025 high school ELA and mathematics assessments in this section of the report. The magnitude of classification consistency and accuracy measures is influenced by several key features of a test's design, including the number of items, the location and number of cut scores, the score distribution, and the reliability and associated SEM. As seen in Table 9.4, classification accuracy conditioned on achievement level ranged from 0.00 to 0.84. As seen in Table 9.5, classification consistency conditioned on achievement level ranged from 0.23 to 0.79. For some mathematics tests, classification accuracy and consistency conditioned on the *Unsatisfactory* level were very low. A possible reason for these relatively low *Unsatisfactory* level values is the fact that there were not enough easy items to distinguish the *Unsatisfactory* level from the *Approaching Basic* performance level.

Table 9.4 Classification Accuracy Conditioned on Level of Achievement

Classification Accuracy							
Administration	**Course**	**Form**	**Unsatisfactory**	**Approaching Basic**	**Basic**	**Mastery**	**Advanced**
Fall 2020	English I	A	0.69	0.68	0.72	0.79	0.68
	English II	A	0.75	0.64	0.71	0.77	0.73
	Algebra I	D	0.22	0.71	0.61	0.82	0.67
	Geometry	D	0.00	0.84	0.68	0.84	0.75
Spring 2021	English I	E	0.76	0.67	0.73	0.80	0.65
	English II	E	0.77	0.62	0.66	0.75	0.75
	Algebra I	E	0.51	0.57	0.59	0.84	0.75
	Geometry	E	0.00	0.84	0.68	0.84	0.75

Table 9.5 Classification Consistency Conditioned on Level of Achievement

Classification Consistency							
Administration	**Course**	**Form**	**Unsatisfactory**	**Approaching Basic**	**Basic**	**Mastery**	**Advanced**
Fall 2020	English I	A	0.62	0.56	0.58	0.70	0.64
	English II	A	0.73	0.52	0.55	0.67	0.69
	Algebra I	D	0.37	0.47	0.49	0.77	0.64
	Geometry	D	0.25	0.59	0.59	0.77	0.73
Spring 2021	English I	E	0.70	0.56	0.59	0.72	0.60
	English II	E	0.75	0.50	0.50	0.63	0.72
	Algebra I	E	0.44	0.45	0.48	0.79	0.72
	Geometry	E	0.23	0.59	0.59	0.78	0.71

Perhaps the most important indices for accountability systems are those for the accuracy and consistency of classification decisions made at specific cut points. To evaluate decisions at specific cut points, the joint distribution of all the performance levels is collapsed into a dichotomized distribution around that specific cut point. As an example, for the LEAP 2025 assessments, a dichotomization at the cut point between the *Basic* and *Mastery* classifications was formed. The proportion of correct classifications below this particular cut point is equal to the sum of all the cells at the *Unsatisfactory*, *Approaching Basic*, and *Basic* levels, and the proportion of correct classifications above this particular cut point is equal to the sum of all the cells at the *Mastery* and *Advanced* levels. Table 9.6 shows the classification accuracy statistics and Table 9.7 shows the classification consistency estimates when conditioned on LEAP 2025 High School cut scores. Table 9.6 shows that classification accuracy at achievement cut points ranged from 0.84 to 0.98. Table 9.7 shows that classification consistency conditioned at achievement cut points ranged from 0.80 to 0.98. Classification consistency and accuracy at achievement cut points tend to be higher values than those conditioned on performance level.

The classification accuracy statistics are at or above 0.84, while the classification consistency statistics are at or above 0.80. These results suggest that consistent and accurate performance-level

classifications are being made for students in Louisiana based on the LEAP 2025 High School assessments.

Table 9.6 Classification Accuracy at Achievement Cut Points

Classification Accuracy						
Administration	**Course**	**Form**	***Unsatisfactory/ Approaching Basic***	***Approaching Basic/ Basic***	***Basic/ Mastery***	***Mastery/ Advanced***
Fall 2020	English I	A	0.94	0.91	0.91	0.96
	English II	A	0.94	0.92	0.91	0.95
	Algebra I	D	0.86	0.85	0.92	0.99
	Geometry	D	0.93	0.87	0.94	0.98
Spring 2021	English I	E	0.95	0.92	0.91	0.96
	English II	E	0.94	0.92	0.91	0.94
	Algebra I	E	0.84	0.87	0.93	0.99
	Geometry	E	0.94	0.87	0.94	0.98

Table 9.7 Classification Consistency at Achievement Cut Points

Classification Consistency						
Administration	**Course**	**Form**	***Unsatisfactory/ Approaching Basic***	***Approaching Basic/ Basic***	***Basic/ Mastery***	***Mastery/ Advanced***
Fall 2020	English I	A	0.92	0.88	0.88	0.94
	English II	A	0.92	0.88	0.88	0.93
	Algebra I	D	0.82	0.80	0.89	0.99
	Geometry	D	0.88	0.82	0.92	0.98
Spring 2021	English I	E	0.92	0.89	0.88	0.94
	English II	E	0.92	0.89	0.87	0.91
	Algebra I	E	0.80	0.82	0.90	0.98
	Geometry	E	0.89	0.82	0.91	0.98

9.7 Convergent Validity

Convergent validity is a subtype of construct validity that can be estimated by the extent to which measures of constructs that theoretically should be related to each other are, in fact, observed as related to each other. Analyses of the internal structure of a test can indicate the extent to which the relationships among test items conform to the construct the test purports to measure. For example, the LEAP 2025 geometry test is designed to measure a single overall construct—geometry achievement; therefore, the items comprising the LEAP 2025 Geometry test should measure only geometry, not language or reading.

This technical report summarizes additional statistics that contribute to construct validity (Cronbach's coefficient alpha is reported previously in this section, and item fit is reported in Chapter 6). The internal consistency coefficient (i.e., Cronbach's alpha) reported is typically measured via correlations among the test items and indicates the degree of the same general construct (Pearson, 2015, page 128). The reliability statistics shown in Table 9.2 are all above 0.87, indicating that items on the 2021 LEAP 2025 High School assessments are homogeneous. For a group of items to be homogeneous, the items must all measure the same construct (i.e., construct validity) or represent the same content domain (i.e., content validity). Because IRT models were used to calibrate test items and to report student scores, item fit is also relevant to construct validity. The extent to which test items function as the IRT model prescribes is relevant to the validation of test scores. In 2019 when the 2021 forms were post-equated, no items were flagged for poor model/data fit.

9.8 Principal Components Analysis

As another measure of construct validity, DRC examined the unidimensionality of each subject-level LEAP 2025 test. One of the underlying assumptions of the IRT models used to scale the LEAP 2025 tests is that the tests being calibrated are unidimensional; that is, items in each subject area measure a single content domain. For example, Algebra I items should measure algebra ability and not reading skills. Standard 1.13 of the *Standards* states:

> If the rationale for a test score interpretation for a given use depends on premises about the relationships among test items or among parts of the test, evidence concerning the internal structure of the test should be provided (26–27).

This section examines the internal structure of the LEAP 2025 tests by evaluating the unidimensionality assumption through principal components analysis (PCA). This analysis seeks evidence that there exists a single primary factor, the first principal component, which accounts for much of the relationship between items. The presence of a single or dominant factor suggests that a test is sufficiently unidimensional (i.e., that it measures one underlying construct).

A PCA was conducted for each subject of the LEAP 2025 assessments. A large first principal component is evident in each analysis. It is common to have additional eigenvalues greater than 1.0, which may suggest the presence of other factors.

For the subjects of the LEAP 2025 assessments, the ratio of variance accounted for by the first factor to variance accounted for by the second factor is sufficiently large to indicate that the unidimensionality assumption holds. All the LEAP 2025 High School tests exhibit first principal components accounting for more than 20% of the test variance (Table 9.8 through Table 9.11), except for the Algebra I spring 2019 administration. To further investigate the unidimensionality of the assessments, the ratio of the first eigenvalue to the second eigenvalue was found and is included in the row below the second component in each table. These ratios show that the first eigenvalue is at least four times as large as the second eigenvalue for all the LEAP 2025 assessments. This substantial difference in magnitude indicates that one factor appears to be dominant and that the LEAP 2025 High School tests are essentially unidimensional.

This evidence supports the claim that there is a dominant dimension underlying the items and tasks in each test and that scores from each test represent performance primarily determined by that ability.

Construct-irrelevant variance, such as factual knowledge irrelevant to doing well in a subject, does not appear to create significant nuisance factors.

Table 9.8 Principal Component Analysis: English I

Administration	Form	Components	Eigenvalue	Percentage of Variance Explained	Cumulative Percentage of Variance Explained
Fall 2020	A	First Component	8.43	26.34	26.34
		Second Component	1.25	3.91	30.25
		Ratio (First/Second)	6.73	-	-
Spring 2021	E	First Component	9.02	28.20	28.20
		Second Component	1.20	3.74	31.93
		Ratio (First/Second)	7.55	-	-

Table 9.9 Principal Component Analysis: English II

Administration	Form	Components	Eigenvalue	Percentage of Variance Explained	Cumulative Percentage of Variance Explained
Fall 2020	A	First Component	8.77	27.40	27.40
		Second Component	1.37	4.29	31.69
		Ratio (First/Second)	6.38	-	-
Spring 2021	E	First Component	8.76	27.37	27.37
		Second Component	1.37	4.27	31.63
		Ratio (First/Second)	6.41	-	-

Table 9.10 Principal Component Analysis: Algebra I

Administration	Form	Components	Eigenvalue	Percentage of Variance Explained	Cumulative Percentage of Variance Explained
Fall 2020	D	First Component	7.30	18.72	18.72
		Second Component	1.40	3.59	22.31
		Ratio (First/Second)	5.21	-	-
Spring 2021	E	First Component	9.10	23.32	23.32
		Second Component	1.26	3.22	26.54
		Ratio (First/Second)	7.24	-	-

Table 9.11 Principal Component Analysis: Geometry

Administration	Form	Components	Eigenvalue	Percentage of Variance Explained	Cumulative Percentage of Variance Explained
Fall 2020	D	First Component	10.56	27.09	27.09
		Second Component	1.51	3.87	30.95
		Ratio (First/Second)	7.01	-	-
Spring 2021	E	First Component	10.25	26.27	26.27
		Second Component	1.40	3.60	29.87
		Ratio (First/Second)	7.30	-	-

9.9 Analyses by Reporting Categories and Subcategories

Three sets of analyses were conducted at the reporting category and subcategory levels for ELA and mathematics content in another attempt to assess the construct validity of the LEAP 2025 assessments. First, correlation coefficients that measure the relationship between the category scores and subcategory scores were computed. Second, the reliability of each category and subcategory was computed. Finally, the SEM was computed for each reportable category and subcategory.

9.10 Correlations among Reporting Categories and Subcategories

This section reports the strength of the interrelationships among the reporting categories or subcategories by computing the correlation between them. Table 9.12 through Table 9.19 report the uncorrected Pearson product-moment (PPM) correlation coefficients, the PPM corrected for attenuation (CAPPM). The PPM among the categories and subcategories is presented below the diagonal portion of the matrix, the CAPPM is presented above the diagonal portion of the matrix, and the reliability coefficients used are shown in Table 9.12 through Table 9.19.

The uncorrected PPM in Table 9.12 through Table 9.19 should be interpreted in the context of the reliability coefficient. In general, lower PPM coefficients are expected between variables that are less reliable. In most cases, the PPM coefficients show that performance on one category or subcategory is moderately to strongly related to performance on another category or subcategory within the same grade and content area. The value of the correlation coefficients will be affected by the limited number of items measuring each category or subcategory. Therefore, caution should be used when comparing the PPM coefficients that measure the relationships between categories or subcategories to those that measure the relationships between content areas. A more modest relationship (i.e., smaller correlation coefficients) is expected to be reported between the categories or subcategories as a consequence of the lower number of items measuring each of the reporting categories. The PPM between two category subscores, for example, may be artificially low because of measurement error.

The CAPPM is reported along with the PPM as indicated by Standard 1.21:

> When statistical adjustments, such as those for restriction of range or attenuation, are made, both adjusted and unadjusted coefficients, as well as the specific procedure used, and all statistics used in the adjustment, should be reported. Estimates of the construct-criterion relationship that remove the effects of measurement error on the test should be clearly reported as adjusted estimates (29).

The attenuation of the PPM can be corrected statistically using Spearman's formula:

$$CAPPM = \frac{r_{xy}}{\sqrt{r_{xx} r_{yy}}}, \tag{9.5}$$

where r_{xy} is the PPM between two categories or GLE strands, r_{xx} is the reliability of one of those categories or GLE strands, and r_{yy} is the reliability of the other category or GLE strand.

The English I and English II assessments show moderate relationships between the reading and writing categories, indicating that these two categories measure some different traits. Across all tables, the CAPPM indicates moderate or strong relationships between the subcategories. The CAPPM for reading vocabulary, written expression, and knowledge and use of language are moderate. In some cases, the CAPPM is greater than 1.0. "Disattenuated values greater than 1.00 indicate that measurement error is not randomly distributed" (Schumacker, 1996). The moderate or strong relationships suggested by the CAPPM in Table 9.12 through Table 9.19 are further evidence of the validity of the test construct. Since the overall content area is comprised of the category or subcategory subscores and the content area is expected to measure a single dimension, these subscores are expected to be moderately or highly related.

Table 9.12 Uncorrected Correlation Coefficient (below Diagonal) and Corrected Correlation Coefficient (above Diagonal) among Categories: English I

Administration	Form	No.	Category	N Items	1	2
Fall 2020	A	1	Reading	29	.	0.85
		2	Writing	4	0.76	.
Spring 2021	E	1	Reading	29	.	0.84
		2	Writing	4	0.76	.

Table 9.13 Uncorrected Correlation Coefficient (below Diagonal) and Corrected Correlation Coefficient (above Diagonal) among Subcategories: English I

	Subcategory Uncorrected and Corrected Correlation Coefficients: English I								
Administration	**Form**	**No.**	**Subcategory**	**N Items**	**1**	**2**	**3**	**4**	**5**
Fall 2020	A	1	Reading Literary Text	7	.	0.99	0.93	0.87	0.86
		2	Reading Informational Text	16	0.72	.	0.98	0.92	0.91
		3	Reading Vocabulary	6	0.49	0.57	.	0.79	0.78
		4	Written Expression	2	0.64	0.74	0.46	.	1.14
		5	Written Knowledge & Use of Language	2	0.64	0.74	0.46	0.94	.
Spring 2021	E	1	Reading Literary Text	11	.	0.99	1.02	0.86	0.83
		2	Reading Informational Text	12	0.74	.	1.01	0.96	0.92
		3	Reading Vocabulary	6	0.67	0.66	.	0.88	0.85
		4	Written Expression	2	0.66	0.74	0.59	.	1.18
		5	Written Knowledge & Use of Language	2	0.65	0.72	0.58	0.94	.

Table 9.14 Uncorrected Correlation Coefficient (below Diagonal) and Corrected Correlation Coefficient (above Diagonal) among Categories: English II

Administration	**Form**	**No.**	**Category**	**N Items**	**1**	**2**
Fall 2020	A	1	Reading	29	.	0.85
		2	Writing	4	0.76	.
Spring 2021	E	1	Reading	29	.	0.83
		2	Writing	4	0.75	.

Table 9.15 Uncorrected Correlation Coefficient (below Diagonal) and Corrected Correlation Coefficient (above Diagonal) among Subcategories: English II

	Subcategory Uncorrected and Corrected Correlation Coefficients: English I								
Administration	**Form**	**No.**	**Subcategory**	**N Items**	**1**	**2**	**3**	**4**	**5**
Fall 2020	A	1	Reading Literary Text	6	.	0.98	0.99	0.85	0.85
		2	Reading Informational Text	16	0.66	.	0.98	0.91	0.91
		3	Reading Vocabulary	7	0.58	0.70	.	0.82	0.83
		4	Written Expression	2	0.57	0.75	0.59	.	1.13
		5	Written Knowledge & Use of Language	2	0.58	0.75	0.60	0.94	.
Spring 2021	E	1	Reading Literary Text	10	.	0.98	0.98	0.80	0.80
		2	Reading Informational Text	11	0.72	.	1.00	0.94	0.92
		3	Reading Vocabulary	8	0.66	0.69	.	0.83	0.83
		4	Written Expression	2	0.62	0.74	0.60	.	1.14
		5	Written Knowledge & Use of Language	2	0.62	0.73	0.60	0.95	.

Table 9.16 Uncorrected Correlation Coefficient (below Diagonal) and Corrected Correlation Coefficient (above Diagonal) among Categories: Algebra I

Administration	Form	No.	Category	N Items	1	2	3	4
Fall 2020	D	1	Major Content	22	.	0.99	0.98	0.96
		2	Additional & Supporting Con	10	0.64	.	1.01	0.86
		3	Expressing Mathematical Reasoning	3	0.64	0.58	.	0.93
		4	Modeling & Application	4	0.69	0.55	0.59	.
Spring 2021	E	1	Major Content	22	.	1.01	0.92	0.97
		2	Additional & Supporting Con	10	0.73	.	0.97	0.97
		3	Expressing Mathematical Reasoning	3	0.67	0.65	.	0.90
		4	Modeling & Application	4	0.76	0.70	0.65	.

Table 9.17 Uncorrected Correlation Coefficient (below Diagonal) and Corrected Correlation Coefficient (above Diagonal) among Categories: Geometry

Administration	Form	No.	Category	N Items	1	2	3	4
Fall 2020	D	1	Major Content	19	.	0.98	1.03	0.95
		2	Additional & Supporting Con	13	0.74	.	1.02	0.98
		3	Expressing Mathematical Reasoning	3	0.76	0.68	.	1.10
		4	Modeling & Application	4	0.74	0.68	0.76	.
Spring 2021	E	1	Major Content	19	.	1.00	0.98	1.00
		2	Additional & Supporting Con	13	0.75	.	0.99	1.03
		3	Expressing Mathematical Reasoning	3	0.76	0.71	.	1.07
		4	Modeling & Application	4	0.75	0.71	0.76	.

Table 9.18 Uncorrected Correlation Coefficient (below Diagonal) and Corrected Correlation Coefficient (above Diagonal) among Subcategories: Algebra I

Administration	Form	No.	Subcategory	N Items	1	2	3
Fall 2020	D	1	A1	6	.	1.11	1.02
		2	A2	7	0.51	.	1.25
		3	A3	6	0.47	0.51	.
Spring 2021	E	1	A1	7	.	1.04	1.16
		2	A2	6	0.54	.	1.13
		3	A3	6	0.60	0.51	.

Table 9.19 Uncorrected Correlation Coefficient (below Diagonal) and Corrected Correlation Coefficient (above Diagonal) among Categories: Geometry

Administration	Form	No.	Subcategory	N Items	1	2
Fall 2020	D	1	A1	11	.	0.98
		2	A2	8	0.71	.
Spring 2021	E	1	A1	11	.	0.96
		2	A2	8	0.64	.

9.11 Reliability of Reporting Categories, or Subcategories

Raw score summary statistics (i.e., mean and standard deviation), Cronbach's (1951) coefficient alpha, and SEM were computed for each of the categories or subcategories by subject using the census data. These statistics are presented in Tables 9.18 through 9.22. Reliability indices, such as Cronbach's coefficient alpha (and resulting SEM), are a function of the number of items on a test, the average covariance between item pairs, and the variance of a test's total score. In general, it is expected that the coefficient alpha would be lower for a category or subcategory assessed by a small number of items than for a category or subcategory assessed by a larger number of items.

9.12 Standard Error of Measurement of Reporting Categories or Subcategories

This chapter also reports the SEM associated with each of the categories and subcategories in Table9.20 through Table 9.27. These SEMs are reported in the raw score metric.

Table 9.20 Mean, Standard Deviation, and Standard Error of Measurement (SEM) of English I Categories

Administration	Form	Category	Number of Items	Number of Score Points	Mean Raw Score	Raw Score Std. Dev.	SEM	Cronbach's Alpha
Fall 2020	A	Reading	29	60	25.11	11.45	4.22	0.86
		Writing	4	30	10.92	7.45	1.99	0.93
Spring 2021	E	Reading	29	60	27.52	12.20	4.25	0.88
		Writing	4	30	9.19	6.83	1.93	0.92

Table 9.21 Mean, Standard Deviation, and Standard Error of Measurement (SEM) of English I Subcategories

Admin.	Form	Subcategory	Number of Items	Number of Score Points	Mean Raw Score	Raw Score Std. Dev.	SEM	Cronbach's Alpha
Fall 2020	A	Reading Literary Text	7	14	5.93	3.38	1.95	0.67
		Reading Informational Text	16	34	14.63	7.08	3.16	0.80
		Reading Vocabulary	6	12	4.56	2.50	1.90	0.42
		Written Expression	2	24	8.29	5.77	2.48	0.82
		Knowledge & Use of Language	2	6	2.63	1.76	0.71	0.83
Spring 2021	E	Reading Literary Text	11	22	10.71	5.00	2.49	0.75
		Reading Informational Text	12	26	10.68	5.52	2.74	0.75
		Reading Vocabulary	6	12	6.12	3.07	2.02	0.57
		Written Expression	2	24	6.87	5.24	2.42	0.79
		Knowledge & Use of Language	2	6	2.32	1.67	0.73	0.81

Table 9.22 Mean, Standard Deviation, and Standard Error of Measurement (SEM) of English II Categories

Administration	Form	Category	Number of Items	Number of Score Points	Mean Raw Score	Raw Score Std. Dev.	SEM	Cronbach's Alpha
Fall 2020	A	Reading	29	60	27.61	10.87	3.85	0.87
		Writing	4	30	11.30	7.47	1.95	0.93
Spring 2021	E	Reading	29	60	27.11	11.77	4.17	0.87
		Writing	4	30	12.80	7.63	1.96	0.93

Table 9.23 Mean, Standard Deviation, and Standard Error of Measurement (SEM) of English II Subcategories

Administration	Form	Subcategory	Number of Items	Number of Score Points	Mean Raw Score	Raw Score Std. Dev.	SEM	Cronbach's Alpha
Fall 2020	A	Reading Literary Text	6	12	5.24	2.53	1.69	0.55
		Reading Informational Text	16	34	13.95	6.59	2.86	0.81
		Reading Vocabulary	7	14	8.42	3.05	1.86	0.63
		Written Expression	2	24	8.59	5.73	2.39	0.83
		Knowledge & Use of Language	2	6	2.71	1.82	0.74	0.84
Spring 2021	E	Reading Literary Text	10	20	9.41	4.62	2.45	0.72
		Reading Informational Text	11	24	9.76	5.05	2.52	0.75
		Reading Vocabulary	8	16	7.94	3.51	2.13	0.63
		Written Expression	2	24	9.69	5.89	2.44	0.83
		Knowledge & Use of Language	2	6	3.11	1.81	0.73	0.84

Table 9.24 Mean, Standard Deviation, and Standard Error of Measurement (SEM) of Algebra I Categories

Administration	Form	Category	Number of Items	Number of Score Points	Mean Raw Score	Raw Score Std. Dev.	SEM	Cronbach's Alpha
Fall 2020	D	Major Content	22	28	10.47	4.69	2.43	0.73
		Additional & Supporting Content	10	14	4.25	2.51	1.64	0.57
		Expressing Mathematical Reasoning	3	11	1.40	1.94	1.26	0.57
		Modeling & Application	4	15	2.69	2.55	1.38	0.71
Spring 2021	E	Major Content	22	28	10.78	5.19	2.40	0.79
		Additional & Supporting Content	10	14	4.65	2.74	1.58	0.67
		Expressing Mathematical Reasoning	3	11	1.53	1.97	1.13	0.67
		Modeling & Application	4	15	2.45	3.07	1.43	0.78

Table 9.25 Mean, Standard Deviation, and Standard Error of Measurement (SEM) of Algebra I Subcategories

Administration	Form	Subcategories	Number of Items	Number of Score Points	Mean Raw Score	Raw Score Std. Dev.	SEM	Cronbach's Alpha
Fall 2020	D	A1	6	7	2.02	1.58	1.09	0.52
		A2	7	12	5.02	2.27	1.75	0.41
		A3	6	6	2.28	1.34	1.03	0.41
Spring 2021	E	A1	7	9	3.43	2.14	1.36	0.60
		A2	6	7	2.27	1.58	1.16	0.46
		A3	6	9	4.13	1.94	1.44	0.45

Table 9.26 Mean, Standard Deviation, and Standard Error of Measurement (SEM) of Geometry Categories

Administration	Form	Category	Number of Items	Number of Score Points	Mean Raw Score	Raw Score Std. Dev.	SEM	Cronbach's Alpha
Fall 2020	D	Major Content	19	26	9.82	5.53	2.26	0.83
		Additional & Supporting Content	13	16	5.86	3.02	1.73	0.67
		Expressing Mathematical Reasoning	3	11	1.59	2.28	1.34	0.65
		Modeling & Application	4	15	1.56	2.54	1.34	0.72
Spring 2021	E	Major Content	19	26	9.97	5.09	2.20	0.81
		Additional & Supporting Content	13	16	5.67	3.08	1.70	0.70
		Expressing Mathematical Reasoning	3	11	1.80	2.49	1.26	0.74
		Modeling & Application	4	15	1.72	2.67	1.49	0.69

Table 9.27 Mean, Standard Deviation, and Standard Error of Measurement (SEM) of Geometry Subcategories

Administration	Form	Subcategories	Number of Items	Number of Score Points	Mean Raw Score	Raw Score Std. Dev.	SEM	Cronbach's Alpha
Fall 2020	D	A1	11	17	7.09	3.76	1.89	0.75
		A2	8	9	2.73	2.19	1.22	0.69
Spring 2021	E	A1	11	17	7.10	3.60	1.77	0.76
		A2	8	9	2.86	1.97	1.27	0.59

9.13 Divergent (Discriminant) Validity

Measures of different constructs should not be highly correlated with each other. Divergent validity is a subtype of construct validity that can be assessed by the extent to which measures of constructs that theoretically should not be related to each other are, in fact, observed as not related to each other. Typically, correlation coefficients among measures of unrelated or distantly related constructs are examined in support of divergent validity.

To assess the divergent validity of the LEAP 2025 High School assessments, correlations were computed between the English I, English II, Algebra I and Geometry total scores for students who took more than one subject test in 2021. These correlations are based on the census data, and the results are shown in Table 9.28 and Table 9.29 for the fall 2020 and spring 2021 administrations respectively. The correlation coefficients ranged from 0.63 to 0.88. The lowest correlation was observed between English II and Algebra I in the fall 2020 administration, and the highest correlation was between English I and English II in the spring 2021 administration. Similar patterns were observed in both administrations. The correlation coefficients suggest that individual student scores across subjects are moderately related,

indicating that these tests measure a similar knowledge base or general underlying ability while still measuring some different traits as planned.

Table 9.28 Inter-Correlation of HS Content Area Scale Scores in Fall Administration

	English I	English II	Algebra I	Geometry
English I	-	0.88	0.75	0.75
English II	0.88 (88)*	-	0.76	0.69
Algebra I	0.75 (1414)	0.76 (325)	-	0.81
Geometry	0.75 (260)	0.69 (1,434)	0.81 (37)	-

*The count of observations in the analysis is in parenthesis

Table 9.29 Inter-Correlation of HS Content Area Scale Scores in Spring Administration

	English I	English II	Algebra I	Geometry
English I	-	0.83	0.72	0.71
English II	0.83 (590)*	-	0.63	0.66
Algebra I	0.72 (34,062)	0.63 (3,792)	-	0.81
Geometry	0.71 (4,747)	0.66 (21,458)	0.81 (429)	-

*The count of observations in the analysis is in parenthesis

9.14 Summary

In summary, the overall purpose of establishing construct validity is to ensure that the interpretation of test scores is supported. Evidence of validity is necessary to justify the use of the LEAP 2025 test scores. This evidence addresses multiple best practices of the testing industry but particularly relates to the following standards.

> *Standard 1.13* If the rationale for a test score interpretation for a given use depends on premises about the relationships among test items or among parts of the test, evidence concerning the internal structure of the test should be provided (26).
>
> *Standard 1.21* When statistical adjustments, such as those for restriction of range or attenuation, are made, both adjusted and unadjusted coefficients, as well as the specific procedure used, and all statistics used in the adjustment, should be reported. Estimates of the construct-criterion relationship that remove the effects of measurement error on the test should be clearly reported as adjusted estimates (29).
>
> *Standard 2.0* Appropriate evidence of reliability/precision should be provided for the interpretation for each intended score use (42).
>
> *Standard 2.3* For each total score, subscore, or combination of scores that is to be interpreted, estimates of relevant indices of reliability/precision should be reported (43).
>
> *Standard 2.13* The standard error of measurement, both overall and conditional (if reported), should be provided in units of each reported score (45).
>
> *Standard 2.14* When possible and appropriate, conditional standard errors of measurement should be reported at several score levels unless there is evidence that the standard error is

> constant across score levels. Where cut scores are specified for selection or classification, the standard errors of measurement should be reported in the vicinity of each cut score (46).
>
> *Standard 2.16* When a test or combination of measures is used to make classification decisions, estimates should be provided of the percentage of test takers who would be classified in the same way on two replications of the procedure (46).
>
> *Standard 2.19* Each method of quantifying the reliability/precision of scores should be described clearly and expressed in terms of statistics appropriate to the method. The sampling procedures used to select test takers for reliability/precision analyses and the descriptive statistics on these samples, subject to privacy obligations where applicable, should be reported (47).

Chapter 10: Fairness

As noted in the *Standards for Educational and Psychological Testing* (American Educational Research Association [AERA], American Psychological Association [APA], & National Council on Measurement in Education [NCME], 2014), there are varying definitions of fairness. This chapter examines fairness as it relates to minimizing bias on a test. This chapter also discusses test performance among varying subgroups assessed by LEAP 2025 assessments. It should be noted that having differences in test performance among subgroups does not mean that a test is unfair—it simply means that groups perform differently on a test. Even when a test is carefully and properly constructed, differences may exist among subgroups as a result of differences in curriculum or learning by students in the subgroups.

This chapter demonstrates how the Leap 2025 assessments adhere to AERA, APA, & NCME Standards 3.1–3.6 and 3.16. These standards are from Chapter 3 of the *Standards*, which is titled "Fairness in Testing." Each of these standards is presented in this chapter.

Standard 3.6 states:

> Where credible evidence indicates that test scores may differ in meaning for relevant subgroups in the intended examinee population, test developers and/or users are responsible for examining the evidence for validity of score interpretations for intended uses for individuals from those subgroups. What constitutes a significant difference in subgroup scores and what actions are taken in response to such differences may be defined by applicable laws (65).

Test scores of examinee subgroups that differ in meaning are an ongoing concern in any large-scale testing program. To lessen the possibility of differences in test score meaning, DRC follows several steps in the item development and item selection processes, as is explained in Section 10.1 of this chapter. In addition, LDOE assessment research and development experts and Louisiana educators conduct content and bias reviews on items during the selection process, as explained in Chapter 3. These practices adhere to Standard 3.3, which states,

> Those responsible for test development should include relevant subgroups in validity, reliability/precision, and other preliminary studies used when constructing the test (64).

The PARCC consortium conducted differential item functioning (DIF) studies of their items prior to PARCC operational administrations. Items are typically evaluated for possible DIF in the field test phase of the test

development process, and any items flagged for DIF are further examined to determine possible bias. During the ELA and mathematics test development process, DRC content experts tried to avoid including PARCC operational items flagged for DIF. Section 10.2 of this chapter explains the steps taken to evaluate LEAP 2025 items through the use of DIF to adhere to Standard 3.3.

In addition, the standardized test administration practices and the extensive training process for test score interpretation for LEAP 2025 comply with Standards 3.4 and 3.5, which state:

> *Standard 3.4* Test takers should receive comparable treatment during the test administration and scoring process (65).
>
> *Standard 3.5* Test developers should specify and document provisions that have been made to test administration and scoring procedures to remove construct-irrelevant barriers for all relevant subgroups in the test-taker population (65).

Section 10.1 of this chapter is also directly relevant to Standards 3.1 and 3.2.

> *Standard 3.1* Those responsible for test development, revision, and administration should design all steps of the testing process to promote valid score interpretations for intended score uses for the widest possible range of individuals and relevant subgroups in the intended population (63).
>
> *Standard 3.2* Test developers are responsible for developing tests that measure the intended construct and for minimizing the potential for tests' being affected by construct-irrelevant characteristics, such as linguistic, communicative, cognitive, cultural, physical, or other characteristics (64).

This chapter explains the steps taken by DRC to minimize words, phrases, and content that may be regarded as offensive by members of particular demographic subgroups. Section 3.2 of Chapter 3 discusses the content and bias review conducted for LEAP 2025. This review is also critical in fulfilling Standards 3.1 and 3.2. In addition to the Louisiana-developed items, the New Meridian operational items used in the 2019 LEAP 2025 forms were critical to the forms construction process. Refer to the New Meridian website for the bias and sensitivity guidelines used and the processes and procedures followed by New Meridian pertaining to these items (see https://newmeridiancorp.org/).

The DIF and reliability analyses in this section are based on the CIA data described in Chapter 6. The impact analyses (scale score mean and standard deviation) are based on the technical report sample described in Chapter 7.

10.1 Minimizing Bias through Careful Test Development

The construction of a test that is fair for all examinees begins in the early stages of planning and development. The item and test development processes that were used to minimize bias are summarized below.

First, careful attention was paid to content validity during the item development and item selection processes. Bias can occur only if the test is measuring different things for different groups. The possibility of bias is reduced by eliminating irrelevant skills or knowledge from the items.

Second, item writers and test developers followed New Meridian Fairness and Sensitivity Guidelines for reducing or eliminating bias. DRC test development staff reviewed all items and other testing materials with these guidelines in mind. Internal editorial reviews were conducted by at least three different people: a content editor who directly supervised the item writers, a style editor, and a content supervisor. The final

test was again reviewed by people in these same roles and was also subjected to an independent review by LDOE assessment research and development specialists.

Third, careful attention was given to item statistics throughout the test development process. As part of the test assembly process, attempts were made to avoid using or reusing items with poor statistical fit or distractors with positive point biserial correlations, since these conditions may indicate that an item is testing a construct irrelevant to what being measured. DIF statistics were also examined during test construction. Items that had exhibited significant DIF against one or more subgroups were removed from further consideration unless it was essential to include them to meet content specifications.

10.2 Evaluating Bias through Differential Item Functioning (DIF) Statistics

After administering the test, an empirical approach known as DIF was used to examine the items. The DIF statistics (see Tables 10.1-10.4) indicate the degree to which members of a particular subgroup perform better or worse than expected on each item as compared to the reference group. The DIF procedures used and the results of these analyses are detailed in this section. It should be noted, however, that all items included in LEAP 2025 were thoroughly reviewed for content and bias by LDOE and DRC content experts to ensure the items do not test knowledge or ability irrelevant to the construct the test intends to measure. Therefore, DIF flags do not necessarily indicate that an item is biased; rather, DIF flags indicate that the item functions differently for equally able members of different groups (Camilli & Shepard, 1994). Items are not necessarily suppressed from operational scoring if they are flagged for DIF.

The position of DRC concerning test bias is based on two general propositions. First, students may differ in their background knowledge, cognitive and academic skills, languages, attitudes, and values. To the degree that these differences are large, no one curriculum and no one set of instructional materials will be equally suitable for all. Therefore, no one test will be equally appropriate for all. Furthermore, it is difficult to specify what amount of difference can be called large and to determine how these differences will affect the outcome of a particular test. Second, schools have been assigned the tasks of developing certain basic cognitive skills and supporting development of these skills equitably among all students. Therefore, there is a need for tests that measure the common skills and bodies of knowledge that are expected of all learners. The test publisher's task is to develop assessments that measure these key cognitive skills without introducing extraneous or construct-irrelevant elements into the performances on which the measurement is based. If these tests require that students have culturally specific knowledge and skills not taught in school, differences in performance among students can occur because of differences in student background and out-of-school learning. Such tests are measuring different things for different groups and can be called biased (Camilli & Shepard, 1994; Green, 1975).

To lessen this bias, DRC strives to minimize the role of extraneous elements, thereby increasing the number of students for whom the test is appropriate. As discussed above and in Chapter 3 of this report, careful attention is given during the item development, test development and test construction processes to lessen the influence of these elements for large numbers of students. Unfortunately, these elements may continue to play a substantial role in some cases. To assess the extent to which items may be performing differently for various subgroups of interest, DIF analyses are conducted after each operational test administration.

DIF statistics are used to quantify differences in item performance between two groups after controlling for examinees' overall achievement level. Two DIF statistics that are commonly used for this purpose are the Mantel-Haenszel (MH) statistic (1959) and the standardized mean difference (SMD) between the reference and focal groups, proposed by Dorans and Schmitt (1991).

The MH statistic is computed as follows (Zwick, Donoghue, & Grima, 1993):

$$\text{Mantel } \chi^2 = \frac{\left(\sum_k F_k - \sum_k E(F_k)\right)^2}{\sum_k \text{Var}(F_k)},$$

where F_k is the sum of scores for the focal group at the *k*th level of the matching variable. Note that the MH statistic is sensitive to N such that larger sample sizes increase the value of chi-square.

In addition to the MH chi-square statistic, the delta statistic (MH-D DIF) was computed for all items. Educational Testing Service (ETS) first developed the MH-D DIF statistic. To compute delta, alpha (the odds ratio) is first computed as follows:

$$\alpha_{MH} = \frac{\sum_{k=1}^{K} N_{r1k} N_{f0k} / N_k}{\sum_{k=1}^{K} N_{f1k} N_{r0k} / N_k},$$

where N_{r1k} is the number of correct responses in the reference group at ability level *k*, N_{f0k} is the number of incorrect responses in the focal group at ability level *k*, N_k is the total number of responses, N_{f1k} is the number of correct responses in the focal group at ability level *k*, and N_{r0k} is the number of incorrect responses in the reference group at ability level *k*. MH-D DIF is then computed as follows:

$$\text{MH-D DIF} = -2.35 \ln(\alpha_{MH})$$

For selected-response items, the MH (χ^2_{MH}) statistic was used to evaluate potential DIF items. In the MH procedure, subgroups are matched by their raw total test score, using a contingency table with *K* ability levels. When applying the MH procedure, the log-odds ratio α is assumed to be constant across the *K* matched levels. The χ^2_{MH} , then, estimates a pooled common-odds ratio. Taking the natural logarithm of the common-odds ratio and its confidence limits and multiplying these with the constant –2.35 may then allow the resulting values to be placed on the MH delta metric (Δ_{MH}) for interpretive purposes. Items were flagged for DIF using the following criteria:

1 Moderate DIF: Significant MH chi-square statistic ($p < 0.05$) and $1.0 \leq |\text{MH D-DIF}| < 1.5$
2 Large DIF: Significant MH chi-square statistic ($p < 0.05$) and $|\text{MH D-DIF}| \geq 1.5$

For constructed-response items, an effect size (ES) statistic based on the MH chi-square will be used. The ES is obtained by dividing the SMD statistics by the standard deviation of the item. The SMD is an effect size index of DIF, which is relatively easy to interpret. The SMD compares the mean of the reference and focal

group, adjusting for the distribution of reference and focal group members on the conditioning variable, which, for these analyses, is the LEAP 2025 raw score. The SMD is computed as follows (Zwick et al., 1993):

$$SMD = p_{Fk}(\sum_{k} m_{Fk} - \sum_{k} m_{Rk}),$$

where p_{Fk} = the proportion of the focal group members at the kth level of the matching variable, $m_{Fk} = 1/N_{F1k}$, and $m_{Rk} = 1/N_{R1k}$. Items are flagged using the same rules that are used in NAEP:

- Moderate DIF: If the MH statistic is significant, ($p < .05$) and $|ES|$ is between 0.17 and 0.25.
- Large DIF: If the MH statistic is significant, ($p < .05$) and $|ES| \geq 0.25$.

A positive DIF value indicates that the item favors the focal group, while a negative value indicates that the item disadvantages the focal group.

DIF Statistics for Demographic Groups

DIF analyses were conducted for groups defined by demographic characteristics. Tables 10.1 to Table 10.4 show the DIF results for the following subgroups:

Gender: Focal group is females; reference group is males.

Ethnicity: Focal groups are Hispanic/Latino, American Indian or Alaska Native, Asian, Black or African American, and two or more races; reference group is white.

Education Classification: Focal group is students who are classified as special education; reference group is all others.

English Learner Status: Focal group is students who are classified as EL; reference group is all others.

Economic Status: Focal group is students who are classified as economically disadvantaged; reference group is all others.

Section 504 Status: Focal group is students who are classified as Section 504; reference group is all others.

Homeless Status: Focal group is students who are classified as homeless; reference group is all others.

Military Affiliation: Focus group is students who are affiliated with the military; reference group is all others.

Foster Care Status: Focus group is students who are in foster care; reference group is all others.

A negative SMD value implies that the focal group has a lower mean item score than the reference group, whereas a positive value implies that the focal group has a higher mean item score than the reference group, conditioned on the matching test score.

The minimum case count for the focal group was set at 200, and the minimum case count for the reference group was set at 400. The DIF analyses are not performed for subgroups of less than 200. In these cases, the statistical procedures do not have sufficient power to detect potential differences.

DIF statistics are produced and examined for all newly field-tested items and for all items being administered for the first time operationally in Louisiana. Since the items on 2021 do not include items in those categories, DIF was not applied to the spring 2021 forms.

DIF Statistics for Test Language

All items on one CBT form of the mathematics test are transadapted from English into Spanish. Transadaptation takes into consideration linguistic and cultural differences and grade-level appropriate words. By accounting for these differences, the achievement of Spanish speakers can be measured in the same way as the achievement of English speakers. Please refer to Appendix B for more information about the transadaptation of Spanish mathematics forms. To help confirm that the test items can be measured similarly regardless of the language in which the items are published, a DIF set of analyses was performed in 2019, when the 2021 forms were initially developed and administered. Two DIF analyses were performed using the 2019 LEAP 2025 mathematics operational items regardless of student count in the reference or focal group. Smaller counts for the groups needed to be tolerated since the overall count for those being administered the Spanish form was low. For these DIF analyses, the reference group was those who were administered the English version of the test and the focal group was those who were administered the Spanish version of the test.

For the first analysis, student responses for the shared operational items between 2018 and 2019 LEAP 2025 mathematics were combined. This approach increased the number of students who took the Spanish versions of the items. The Mantel-Haenszel (MH) and the Standardized Mean Difference (SMD) DIF procedures were performed on these common items. The second analysis focused on the items that were not shared between the 2018 and 2019 administrations. Although the MH and the SMD DIF procedures were performed on all 2019 LEAP 2025 operational items, the DIF flags were applied, where appropriate, to items that were not shared between 2018 and 2019.

For both analyses, DIF results were carefully reviewed whenever sample sizes were smaller than the required minimum sample size and when an item showed large (i.e., C) DIF. Table 10.1 summarizes how many items overall exhibited moderate or large DIF in mathematics.

Table 10.1 2019 LEAP 2025 DIF Statistics: Number of Flagged Items, Mathematics

DIF Statistics: Mathematics				Count of Items at DIF Magnitude			
				Moderate		Large	
Content Area	Number of Items	Category	Group	B-	B+	C-	C+
Algebra I	14	Test Language	Spanish		2	4	
Geometry	16	Test Language	Spanish		1	1	

10.3 Evaluating Bias through Impact Analysis

The impact of achievement testing on subgroups can be determined and reported in the form of average scores and also in terms of test score reliability.

Table 10.6 through Table 10.13 present the number of students and test form reliability statistics (i.e., coefficient alpha; see Chapter 9). Scale score means, standard deviations, and effect sizes (i.e., Cohen's *d*) for the various subgroups of interest are reported by form in Table 10.14 though Table 10.21

10.4 Reliability

Tables 10.2–10.9 show the test form reliability coefficients and SEM by student gender, ethnicity, education classification, economic status, EL status, migrant status, Section 504 status, homeless status, military affiliation, and foster care status. The reliability coefficients for English I and II forms ranged from 0.81 to 0.93 and from 0.87 to 0.93 for the fall 2020 and spring 2021 administrations, respectively. For algebra I and geometry, the reliability coefficients ranged from 0.67 to 0.91 and from 0.81 to 0.94 for the fall 2020 and spring 2021 administrations, respectively. These analyses show that the test reliability is of acceptable magnitude for all the subgroups. Note that the reliability coefficients are based on initial testers and are NR for subgroups with fewer than 10 students.

Table 10.2 Fall 2020 Administration English I Reliability and SEM by Subgroup

Group	N Count	Cronbach's Alpha	SEM
All Students	≥5,740	0.90	5.55
Gender			
Female	≥2,800	0.90	5.58
Male	≥2,930	0.90	5.41
Ethnicity			
Hispanic/Latino	≥540	0.91	5.45
American Indian or Alaska Native	≥20	0.84	5.76
Asian	≥100	0.91	5.39
Black or African American	≥2,600	0.88	5.41
Native Hawaiian or Other Pacific	<10	NR	NR
White	≥2,310	0.89	5.68
Two or More Races	≥140	0.87	5.6
Education Classification			
Regular Education	≥5,320	0.90	5.57
Special Education	≥270	0.85	4.54
Gifted or Talented	≥140	0.86	5.32
Economic Status*			
Economically Disadvantaged	<10	NR	NR
Not Economically Disadvantaged	≥5,000	0.90	5.56
English Learner Status			
Not English Learner	≥5,460	0.90	5.57
English Learner	≥270	0.85	4.85
Migrant Status			
Migrant	<10	NR	NR
Not Migrant	≥5,730	0.90	5.55
Section 504 Status			
Non-Section 504	≥5,190	0.90	5.57
Section 504	≥540	0.88	5.27
Homeless Status			
Not Homeless	≥5,690	0.90	5.55
Homeless	≥40	0.89	5.07
Military Affiliation			
Not Military Affiliated	≥5,700	0.90	5.55
Military Affiliated	≥40	0.81	6.34
Foster Care Status			
Not in Foster Care	≥5,730	0.90	5.55
Foster Care	<10	NR	NR

*Economic Status was not available for all students.

Table 10.3 Spring 2021 Administration English I Reliability and SEM by Subgroup

Group	Form E		
	N Count	Cronbach's Alpha	SEM
All Students	≥45,840	0.91	5.39
Gender			
Female	≥22,350	0.91	5.45
Male	≥23,490	0.91	5.27
Ethnicity			
Hispanic/Latino	≥3,890	0.93	5.14
American Indian or Alaska Native	≥320	0.90	5.47
Asian	≥790	0.92	5.38
Black or African American	≥19,000	0.89	5.25
Native Hawaiian or Other Pacific	≥30	0.89	5.40
White	≥20,590	0.90	5.48
Two or More Races	≥1,190	0.89	5.51
Education Classification			
Regular Education	≥38,820	0.90	5.42
Special Education	≥ 4,370	0.87	4.68
Gifted or Talented	≥ 2,640	0.87	5.35
Economic Status*			
Economically Disadvantaged	≥29,520	0.90	5.31
Not Economically Disadvantaged	≥13,090	0.89	5.48
English Learner Status			
Not English Learner	≥44,060	0.91	5.42
English Learner	≥1,780	0.87	4.48
Migrant Status			
Migrant	≥70	0.91	5.28
Not Migrant	≥45,770	0.91	5.39
Section 504 Status			
Non-Section 504	≥41,640	0.91	5.40
Section 504	≥4,200	0.89	5.17
Homeless Status			
Not Homeless	≥45,230	0.91	5.39
Homeless	≥600	0.90	5.24
Military Affiliation			
Not Military Affiliated	≥45,210	0.91	5.39
Military Affiliated	≥620	0.89	5.50
Foster Care Status			
Not in Foster Care	≥45,680	0.91	5.39
Foster Care	≥150	0.92	5.15

*Economic Status was not available for all students.

Table 10.4 Fall 2020 Administration English II Reliability and SEM by Subgroup

Group	N Count	Cronbach's Alpha	SEM
All Students	≥8,330	0.91	5.17
Gender			
Female	≥4,220	0.91	5.17
Male	≥4,110	0.91	5.07
Ethnicity			
Hispanic/Latino	≥980	0.92	5.08
American Indian or Alaska Native	≥40	0.90	5.02
Asian	≥210	0.93	5.33
Black or African American	≥3,540	0.89	5.04
Native Hawaiian or Other Pacific	<10	NR	NR
White	≥3,350	0.90	5.19
Two or More Races	≥180	0.90	5.16
Education Classification			
Regular Education	≥7,660	0.90	5.16
Special Education	≥320	0.86	4.69
Gifted or Talented	≥340	0.88	5.06
Economic Status*			
Economically Disadvantaged	<10	NR	NR
Not Economically Disadvantaged	≥7,440	0.91	5.17
English Learner Status			
Not English Learner	≥7,830	0.91	5.18
English Learner	≥500	0.83	4.61
Migrant Status			
Migrant	<10	NR	NR
Not Migrant	≥8,320	0.91	5.17
Section 504 Status			
Non-Section 504	≥7,590	0.91	5.18
Section 504	≥740	0.90	5.04
Homeless Status			
Not Homeless	≥8,270	0.91	5.17
Homeless	≥50	0.88	5.19
Military Affiliation			
Not Military Affiliated	≥8,300	0.91	5.17
Military Affiliated	≥20	0.90	5.40
Foster Care Status			
Not in Foster Care	≥8,320	0.91	5.17
Foster Care	<10	NR	NR

*Economic Status was not available for all students.

Table 10.5 Spring 2021 Administration English II Reliability and SEM by Subgroup

Group	Form E		
	N Count	Cronbach's Alpha	SEM
All Students	≥39,950	0.91	5.5
Gender			
Female	≥19,930	0.90	5.49
Male	≥20,010	0.91	5.42
Ethnicity			
Hispanic/Latino	≥2,760	0.92	5.49
American Indian or Alaska Native	≥280	0.90	5.53
Asian	≥650	0.92	5.43
Black or African American	≥16,520	0.89	5.41
Native Hawaiian or Other Pacific	≥30	0.91	5.61
White	≥18,750	0.89	5.52
Two or More Races	≥940	0.89	5.57
Education Classification			
Regular Education	≥34,200	0.90	5.51
Special Education	≥3,420	0.87	5.03
Gifted or Talented	≥2,330	0.87	5.30
Economic Status*			
Economically Disadvantaged	≥24,930	0.90	5.48
Not Economically Disadvantaged	≥12,760	0.89	5.46
English Learner Status			
Not English Learner	≥38,880	0.91	5.51
English Learner	≥1,060	0.84	5.06
Migrant Status			
Migrant	≥40	0.92	5.28
Not Migrant	≥39,910	0.91	5.50
Section 504 Status			
Non-Section 504	≥36,360	0.91	5.50
Section 504	≥3,580	0.90	5.46
Homeless Status			
Not Homeless	≥39,450	0.91	5.50
Homeless	≥490	0.90	5.41
Military Affiliation			
Not Military Affiliated	≥39,450	0.91	5.50
Military Affiliated	≥500	0.90	5.48
Foster Care Status			
Not in Foster Care	≥39,820	0.91	5.50
Foster Care	≥120	0.90	5.46

*Economic Status was not available for all students.

Table 10.6 Fall 2020 Administration Algebra I Reliability and SEM by Subgroup

Group	N Count	Cronbach's Alpha	SEM
All Students	≥4,290	0.87	3.58
Gender			
Female	≥2,080	0.88	3.61
Male	≥2,210	0.87	3.54
Ethnicity			
Hispanic/Latino	≥410	0.86	3.47
American Indian or Alaska Native	≥10	0.90	3.86
Asian	≥90	0.91	4.09
Black or African American	≥1,950	0.82	3.29
Native Hawaiian or Other Pacific	<10	NR	NR
White	≥1,720	0.87	3.79
Two or More Races	≥80	0.86	3.60
Education Classification			
Regular Education	≥4,030	0.87	3.59
Special Education	≥180	0.73	2.91
Gifted or Talented	≥70	0.90	3.99
Economic Status*			
Economically Disadvantaged	<10	NR	NR
Not Economically Disadvantaged	≥3,680	0.87	3.61
English Learner Status			
Not English Learner	≥4,090	0.87	3.60
English Learner	≥190	0.78	3.03
Migrant Status			
Migrant	<10	NR	NR
Not Migrant	≥4,290	0.87	3.58
Section 504 Status			
Non-Section 504	≥3,870	0.87	3.60
Section 504	≥420	0.86	3.32
Homeless Status			
Not Homeless	≥4,270	0.87	3.58
Homeless	≥20	0.68	3.03
Military Affiliation			
Not Military Affiliated	≥4,270	0.87	3.58
Military Affiliated	≥10	0.80	3.62
Foster Care Status			
Not in Foster Care	≥4,290	0.87	3.58
Foster Care	<10	NR	NR

*Economic Status was not available for all students.

Table 10.7 Spring 2021 Administration Algebra I Reliability and SEM by Subgroup

Group	Form E		
	N Count	Cronbach's Alpha	SEM
All Students	≥47,920	0.90	3.56
Gender			
Female	≥23,460	0.90	3.60
Male	≥24,460	0.91	3.52
Ethnicity			
Hispanic/Latino	≥3,820	0.90	3.44
American Indian or Alaska Native	≥320	0.90	3.59
Asian	≥770	0.93	4.08
Black or African American	≥20,460	0.85	3.19
Native Hawaiian or Other Pacific	≥40	0.92	3.69
White	≥21,170	0.91	3.80
Two or More Races	≥1,310	0.91	3.69
Education Classification			
Regular Education	≥40,410	0.89	3.54
Special Education	≥4,600	0.81	2.88
Gifted or Talented	≥2,900	0.92	4.18
Economic Status*			
Economically Disadvantaged	≥31,190	0.87	3.34
Not Economically Disadvantaged	≥13,150	0.91	3.91
English Learner Status			
Not English Learner	≥46,310	0.90	3.58
English Learner	≥1,600	0.82	2.93
Migrant Status			
Migrant	≥60	0.90	3.46
Not Migrant	≥47,850	0.90	3.56
Section 504 Status			
Non-Section 504	≥43,420	0.91	3.59
Section 504	≥4,490	0.87	3.23
Homeless Status			
Not Homeless	≥47,260	0.90	3.57
Homeless	≥650	0.87	3.18
Military Affiliation			
Not Military Affiliated	≥47,270	0.90	3.56
Military Affiliated	≥640	0.92	3.93
Foster Care Status			
Not in Foster Care	≥47,740	0.90	3.56
Foster Care	≥170	0.86	3.18

*Economic Status was not available for all students.

Table 10.8 Fall 2020 Administration Geometry Reliability and SEM by Subgroup

Group	N Count	Cronbach's Alpha	SEM
All Students	≥5,610	0.91	3.52
Gender			
Female	≥2,970	0.91	3.53
Male	≥2,640	0.92	3.51
Ethnicity			
Hispanic/Latino	≥570	0.91	3.47
American Indian or Alaska Native	≥30	0.90	3.58
Asian	≥140	0.93	4.14
Black or African American	≥2,510	0.86	3.12
Native Hawaiian or Other Pacific	<10	NR	NR
White	≥2,220	0.91	3.77
Two or More Races	≥120	0.91	3.81
Education Classification			
Regular Education	≥5,220	0.91	3.49
Special Education	≥130	0.84	2.81
Gifted or Talented	≥260	0.92	4.04
Economic Status*			
Economically Disadvantaged	<10	NR	NR
Not Economically Disadvantaged	≥4,880	0.92	3.58
English Learner Status			
Not English Learner	≥5,390	0.91	3.54
English Learner	≥210	0.87	2.97
Migrant Status			
Migrant	<10	NR	NR
Not Migrant	≥5,610	0.91	3.52
Section 504 Status			
Non-Section 504	≥5,160	0.91	3.54
Section 504	≥440	0.90	3.29
Homeless Status			
Not Homeless	≥5,550	0.91	3.53
Homeless	≥50	0.87	3.20
Military Affiliation			
Not Military Affiliated	≥5,590	0.91	3.52
Military Affiliated	≥10	0.93	3.86
Foster Care Status			
Not in Foster Care	≥5,610	0.91	3.52
Foster Care	<10	NR	NR

*Economic Status was not available for all students.

Table 10.9 Spring 2021 Administration Geometry Reliability and SEM by Subgroup

Group	Form E		
	N Count	Cronbach's Alpha	SEM
All Students	≥34,280	0.91	3.53
Gender			
Female	≥18,210	0.91	3.53
Male	≥16,070	0.92	3.54
Ethnicity			
Hispanic/Latino	≥2,290	0.91	3.48
American Indian or Alaska Native	≥200	0.91	3.62
Asian	≥690	0.93	4.16
Black or African American	≥14,000	0.85	3.07
Native Hawaiian or Other Pacific	≥20	0.94	3.94
White	≥16,270	0.91	3.77
Two or More Races	≥770	0.91	3.62
Education Classification			
Regular Education	≥30,000	0.90	3.49
Special Education	≥1,900	0.85	2.83
Gifted or Talented	≥2,370	0.92	4.11
Economic Status*			
Economically Disadvantaged	≥20,320	0.89	3.28
Not Economically Disadvantaged	≥11,960	0.91	3.87
English Learner Status			
Not English Learner	≥33,550	0.91	3.54
English Learner	≥720	0.86	2.94
Migrant Status			
Migrant	≥20	0.91	3.73
Not Migrant	≥34,260	0.91	3.53
Section 504 Status			
Non-Section 504	≥31,730	0.91	3.55
Section 504	≥2,550	0.90	3.28
Homeless Status			
Not Homeless	≥33,890	0.91	3.54
Homeless	≥390	0.87	3.14
Military Affiliation			
Not Military Affiliated	≥33,770	0.91	3.53
Military Affiliated	≥510	0.92	3.82
Foster Care Status			
Not in Foster Care	≥34,200	0.91	3.54
Foster Care	≥70	0.87	3.01

*Economic Status was not available for all students.

10.5 Effect Size

One way to evaluate the magnitude of the standardized mean difference (SMD) is to calculate the ES. Cohen's *d* was used to calculate the ES and is given by the following formula:

$$d = \frac{\overline{x_a} - \overline{x_b}}{\sqrt{\frac{(n_a - 1)s_a^2 + (n_b - 1)s_b^2}{(n_a + n_b) - 2}}},$$

where $\overline{x_a}$ is the mean score of group A, $\overline{x_b}$ is the mean score of group B, s_a^2 is the variance of group A, s_b^2 is the variance of group B, n_a is the number of students in group A, and n_b is the number of students in group B.

Cohen's *d*, then, expresses the difference in group means in terms of the standard deviation. For example, if *d* = 0.34 for two groups, then it may be interpreted that the SMD between the two groups is 0.34 of the pooled standard deviation. Cohen (1988) offered guidelines for interpreting the meaning of the *d* statistic: *d* = 0.20 is a small ES, *d* = 0.50 is a medium ES, and *d* = 0.80 is a large ES.

Using Cohen's (1988) guidelines, certain trends become apparent in Tables 10.14–10.17. Results are NR for subgroups with fewer than 10 students. If the effect size is negative, that means the group outperforms the group to which it's being compared. For example, in Table 10.10 the effect size for the group female is -0.43 indicating that there is a small difference in performance and females are outperforming males. For all subjects across both the fall and spring administrations, small differences in test scores were seen between females and males, with females slightly outperforming males. Mean scale scores and ESs show that Asian and white students tend to outperform other ethnicity groups across subjects. For most ELA and mathematics tests, there were clear performance differences between regular education and special education students in Education Classification and Not English Learner and English Learner in EL status.

Table 10.10 Fall 2020 Administration Impact Analysis: English I

Group	N	Scale Score Mean	Scale Score Std. Dev.	Effect Size
All Students	≥7,070	732.15	37.04	
Gender				
Male	≥3,840	724.96	35.50	
Female	≥3,230	740.69	37.03	-0.43
Ethnicity				
White	≥2,560	749.48	36.51	
Hispanic/Latino	≥780	719.27	37.89	0.82
American Indian or Alaska Native	≥30	738.61	32.12	0.30
Asian	≥120	751.59	41.32	-0.06
Black or African American	≥3,410	720.79	31.35	0.85
Native Hawaiian or Other Pacific	<10	NR	NR	NR
Two or More Races	≥150	742.57	31.93	0.19
Education Classification				
Regular Education	≥6,350	734.27	35.97	
Special Education	≥570	697.47	23.52	1.05
Gifted or Talented	≥140	776.21	33.76	-1.17
Economic Status				
Not Economically Disadvantaged	—	—	—	
Economically Disadvantaged	—	—	—	—
English Learner Status				
Not English Learner	≥6,540	734.83	36.56	
English Learner	≥530	699.08	25.08	1.00
Migrant Status				
Nonmigrant	≥7,060	732.20	37.04	
Migrant	≥10	709.00	28.25	0.63
Section 504 Status				
Non-Section 504	≥6,290	734.07	37.38	
Section 504	≥780	716.63	29.97	0.48
Homeless Status				
Not Homeless	≥7,000	732.37	37.02	
Homeless	≥70	710.77	32.52	0.58
Military Affiliation				
Not Military Affiliated	≥7,030	732.09	37.07	
Military Affiliated	≥40	740.89	32.19	-0.24
Foster Care Status				
Not in Foster Care	≥7,070	732.17	37.04	
Foster Care	<10	NR	NR	NR

Table 10.11 Spring 2021 Administration Impact Analysis: English I

Group	N	Scale Score Mean	Scale Score Std. Dev.	Effect Size
All Students	≥46,730	739.38	36.37	
Gender				
Male	≥24,180	733.03	36.27	
Female	≥22,540	746.18	35.24	-0.37
Ethnicity				
White	≥20,590	752.42	33.85	
Hispanic/Latino	≥3,970	726.94	40.21	0.73
American Indian or Alaska Native	≥320	742.93	33.61	0.28
Asian	≥790	765.93	38.73	-0.40
Black or African American	≥19,790	726.73	32.58	0.77
Native Hawaiian or Other Pacific	≥30	748.62	31.37	0.11
Two or More Races	≥1,210	746.50	32.89	0.17
Education Classification				
Regular Education	≥39,340	740.79	33.90	
Special Education	≥4,730	704.83	29.29	1.08
Gifted or Talented	≥2,650	780.12	30.25	-1.17
Economic Status				
Not Economically Disadvantaged	≥13,200	759.91	32.61	
Economically Disadvantaged	≥30,430	731.41	34.43	0.84
English Learner Status				
Not English Learner	≥44,920	741.09	35.62	
English Learner	≥1,800	696.65	28.04	1.26
Migrant Status				
Nonmigrant	≥46,650	739.39	36.37	
Migrant	≥70	729.36	35.86	0.28
Section 504 Status				
Non-Section 504	≥42,280	740.88	36.43	
Section 504	≥4,450	725.15	32.57	0.44
Homeless Status				
Not Homeless	≥46,080	739.60	36.35	
Homeless	≥640	723.17	34.45	0.45
Military Affiliation				
Not Military Affiliated	≥46,100	739.13	36.34	
Military Affiliated	≥620	757.28	34.24	-0.50
Foster Care Status				
Not in Foster Care	≥46,570	739.44	36.35	
Foster Care	≥160	722.79	38.76	0.46

Table 10.12 Fall 2020 Administration Impact Analysis: English II

Group	N	Scale Score Mean	Scale Score Std. Dev.	Effect Size
All Students	≥8,880	736.70	42.27	
Gender				
Male	≥4,490	729.79	41.94	
Female	≥4,380	743.77	41.44	-0.34
Ethnicity				
White	≥3,420	754.05	40.44	
Hispanic/Latino	≥1,100	722.07	44.29	0.77
American Indian or Alaska Native	≥40	742.62	38.61	0.28
Asian	≥220	760.90	51.23	-0.17
Black or African American	≥3,880	723.30	35.89	0.81
Native Hawaiian or Other Pacific	<10	NR	NR	NR
Two or More Races	≥190	751.67	39.54	0.06
Education Classification				
Regular Education	≥8,100	736.87	40.52	
Special Education	≥420	693.77	31.47	1.07
Gifted or Talented	≥340	785.88	38.39	-1.21
Economic Status				
Not Economically Disadvantaged	—	—	—	
Economically Disadvantaged	—	—	—	—
English Learner Status				
Not English Learner	≥8,260	740.12	41.10	
English Learner	≥620	691.12	29.17	1.21
Migrant Status				
Nonmigrant	≥8,870	736.70	42.27	
Migrant	<10	NR	NR	NR
Section 504 Status				
Non-Section 504	≥8,060	738.18	42.46	
Section 504	≥810	722.06	37.36	0.38
Homeless Status				
Not Homeless	≥8,810	736.87	42.24	
Homeless	≥60	712.20	38.83	0.58
Military Affiliation				
Not Military Affiliated	≥8,850	736.66	42.26	
Military Affiliated	≥20	748.83	42.88	-0.29
Foster Care Status				
Not in Foster Care	≥8,870	736.70	42.27	
Foster Care	<10	NR	NR	NR

Table 10.13 Spring 2021 Administration Impact Analysis: English II

Group	N	Scale Score Mean	Scale Score Std. Dev.	Effect Size
All Students	≥40,830	743.93	46.18	
Gender				
Male	≥20,650	736.12	46.32	
Female	≥20,180	751.91	44.64	-0.35
Ethnicity				
White	≥18,780	760.37	42.78	
Hispanic/Latino	≥2,970	732.87	49.23	0.63
American Indian or Alaska Native	≥280	747.45	42.45	0.30
Asian	≥660	776.11	50.99	-0.37
Black or African American	≥17,130	725.91	41.61	0.82
Native Hawaiian or Other Pacific	≥30	762.61	48.08	-0.05
Two or More Races	≥950	754.15	42.39	0.15
Education Classification				
Regular Education	≥34,810	745.29	43.27	
Special Education	≥3,670	699.01	36.24	1.09
Gifted or Talented	≥2,340	794.12	39.64	-1.13
Economic Status				
Not Economically Disadvantaged	≥12,880	768.18	42.31	
Economically Disadvantaged	≥25,730	733.27	43.39	0.81
English Learner Status				
Not English Learner	≥39,570	745.57	45.59	
English Learner	≥1,260	692.39	32.94	1.18
Migrant Status				
Nonmigrant	≥40,790	743.94	46.18	
Migrant	≥40	735.89	45.90	0.17
Section 504 Status				
Non-Section 504	≥37,060	745.74	46.15	
Section 504	≥3,770	726.14	42.61	0.43
Homeless Status				
Not Homeless	≥40,310	744.16	46.17	
Homeless	≥510	725.92	43.17	0.40
Military Affiliation				
Not Military Affiliated	≥40,330	743.69	46.15	
Military Affiliated	≥500	762.91	45.04	-0.42
Foster Care Status				
Not in Foster Care	≥40,700	744.00	46.17	
Foster Care	≥40	739.00	35.47	0.22

Table 10.14 Fall 2020 Administration Impact Analysis: Algebra I

Group	N	Scale Score Mean	Scale Score Std. Dev.	Effect Size
All Students	≥5,100	728.25	30.97	
Gender				
Male	≥2,670	726.66	30.36	
Female	≥2,430	730.00	31.54	-0.11
Ethnicity				
White	≥1,930	740.70	31.41	
Hispanic/Latino	≥500	722.36	30.50	0.59
American Indian or Alaska Native	≥10	742.22	31.71	-0.05
Asian	≥90	757.89	36.83	-0.54
Black or African American	≥2,440	718.13	25.80	0.79
Native Hawaiian or Other Pacific	<10	NR	NR	NR
Two or More Races	≥90	733.63	29.30	0.23
Education Classification				
Regular Education	≥4,680	729.58	30.52	
Special Education	≥330	703.18	20.87	0.88
Gifted or Talented	≥80	756.07	35.64	-0.87
Economic Status				
Not Economically Disadvantaged	—	—	—	
Economically Disadvantaged	—	—	—	—
English Learner Status				
Not English Learner	≥4,810	729.49	30.93	
English Learner	≥280	707.53	23.45	0.72
Migrant Status				
Nonmigrant	≥5,100	728.26	30.98	
Migrant	<10	NR	NR	NR
Section 504 Status				
Non-Section 504	≥4,560	729.29	31.13	
Section 504	≥540	719.48	28.17	0.32
Homeless Status				
Not Homeless	≥5,080	728.31	30.99	
Homeless	≥20	714.96	23.19	0.43
Military Affiliation				
Not Military Affiliated	≥5,080	728.24	30.99	
Military Affiliated	≥10	730.05	27.24	-0.06
Foster Care Status				
Not in Foster Care	≥5,100	728.27	30.97	
Foster Care	<10	NR	NR	NR

Table 10.15 Spring 2021 Administration Impact Analysis: Algebra I

Group	N	Scale Score Mean	Scale Score Std. Dev.	Effect Size
All Students	≥48,500	733.34	34.40	
Gender				
Male	≥24,780	731.17	35.00	
Female	≥23,720	735.60	33.60	-0.13
Ethnicity				
White	≥21,190	745.86	34.52	
Hispanic/Latino	≥3,880	727.60	33.70	0.53
American Indian or Alaska Native	≥320	734.65	32.88	0.32
Asian	≥760	766.21	40.92	-0.59
Black or African American	≥20,950	720.16	28.01	0.82
Native Hawaiian or Other Pacific	≥40	743.33	35.60	0.07
Two or More Races	≥1,330	738.61	34.98	0.21
Education Classification				
Regular Education	≥40,760	733.71	32.33	
Special Education	≥4,820	706.67	25.49	0.85
Gifted or Talented	≥2,910	772.28	36.08	-1.18
Economic Status				
Not Economically Disadvantaged	≥13,270	752.90	34.79	
Economically Disadvantaged	≥31,960	725.11	30.30	0.88
English Learner Status				
Not English Learner	≥46,910	734.17	34.35	
English Learner	≥1,590	708.83	25.68	0.74
Migrant Status				
Nonmigrant	≥48,430	733.34	34.40	
Migrant	≥60	733.90	32.49	-0.02
Section 504 Status				
Non-Section 504	≥43,850	734.67	34.63	
Section 504	≥4,650	720.81	29.26	0.41
Homeless Status				
Not Homeless	≥47,820	733.55	34.41	
Homeless	≥670	718.49	29.66	0.44
Military Affiliation				
Not Military Affiliated	≥47,850	733.04	34.26	
Military Affiliated	≥650	755.03	37.59	-0.64
Foster Care Status				
Not in Foster Care	≥48,320	733.40	34.40	
Foster Care	≥170	715.97	28.75	0.51

Table 10.16 Fall 2020 Administration Impact Analysis: Geometry

Group	N	Scale Score Mean	Scale Score Std. Dev.	Effect Size
All Students	≥5,680	733.37	26.70	
Gender				
Male	≥2,680	733.92	27.15	
Female	≥3,000	732.89	26.28	0.04
Ethnicity				
White	≥2,220	744.55	25.32	
Hispanic/Latino	≥600	730.68	27.09	0.54
American Indian or Alaska Native	≥30	736.16	22.76	0.33
Asian	≥130	759.74	30.56	-0.59
Black or African American	≥2,550	722.21	22.04	0.95
Native Hawaiian or Other Pacific	<10	NR	NR	NR
Two or More Races	≥120	744.94	25.13	-0.02
Education Classification				
Regular Education	≥5,280	732.53	25.92	
Special Education	≥140	714.96	19.31	0.68
Gifted or Talented	≥260	760.56	28.28	-1.08
Economic Status				
Not Economically Disadvantaged	—	—	—	
Economically Disadvantaged	—	—	—	—
English Learner Status				
Not English Learner	≥5,440	734.17	26.61	
English Learner	≥240	715.63	22.32	0.70
Migrant Status				
Nonmigrant	≥5,680	733.38	26.71	
Migrant	<10	NR	NR	NR
Section 504 Status				
Non-Section 504	≥5,220	733.97	26.78	
Section 504	≥450	726.62	24.75	0.28
Homeless Status				
Not Homeless	≥5,620	733.48	26.71	
Homeless	≥60	723.53	23.46	0.37
Military Affiliation				
Not Military Affiliated	≥5,660	733.33	26.67	
Military Affiliated	≥10	747.17	32.87	-0.52
Foster Care Status				
Not in Foster Care	≥5,680	733.39	26.69	
Foster Care	<10	NR	NR	NR

Table 10.17 Spring 2020 Administration Impact Analysis: Geometry

Group	N	Scale Score Mean	Scale Score Std. Dev.	Effect Size
All Students	≥34,260	734.21	26.60	
Gender				
Male	≥16,020	735.03	27.21	
Female	≥18,230	733.49	26.04	0.06
Ethnicity				
White	≥16,130	744.57	25.48	
Hispanic/Latino	≥2,330	731.29	25.74	0.52
American Indian or Alaska Native	≥200	737.77	25.02	0.27
Asian	≥680	758.18	30.39	-0.53
Black or African American	≥14,070	721.33	21.35	0.98
Native Hawaiian or Other Pacific	≥20	751.62	30.69	-0.28
Two or More Races	≥780	738.17	25.56	0.25
Education Classification				
Regular Education	≥29,950	733.45	25.20	
Special Education	≥1,920	713.93	21.10	0.78
Gifted or Talented	≥2,380	760.17	28.61	-1.05
Economic Status				
Not Economically Disadvantaged	≥12,010	747.65	26.16	
Economically Disadvantaged	≥20,580	726.88	23.80	0.84
English Learner Status				
Not English Learner	≥33,510	734.63	26.57	
English Learner	≥750	715.20	20.61	0.73
Migrant Status				
Nonmigrant	≥34,230	734.21	26.60	
Migrant	≥20	735.95	27.55	-0.07
Section 504 Status				
Non-Section 504	≥31,670	734.81	26.65	
Section 504	≥2,590	726.88	24.95	0.30
Homeless Status				
Not Homeless	≥33,860	734.35	26.62	
Homeless	≥390	721.79	21.89	0.47
Military Affiliation				
Not Military Affiliated	≥33,740	734.03	26.55	
Military Affiliated	≥510	746.17	27.50	-0.46
Foster Care Status				
Not in Foster Care	≥34,180	734.24	26.60	
Foster Care	≥80	720.88	23.78	0.50

Additional data for scale score means are provided in Tables 10.18 and 10.21. These tables report the number of students, scale score means, and standard deviations for each Special Education Classification. Groups that have fewer than 10 students are not reported (NR) in the tables.

Table 10.18 Special Education Classification Scale Score Means and Standard Deviations: English I

Admin.	Form	Group	Yes			No		
			N	Mean	Std. Dev.	N	Mean	Std. Dev.
Fall 2020	B	Gifted	≥60	794.39	25.69	≥7,010	731.60	36.67
		Talented	≥80	763.15	32.91	≥6,990	731.77	36.93
		Autism	<50	NR	NR	≥7,050	732.20	37.03
		Deaf-Blindness	<50	NR	NR	≥7,070	732.15	37.04
		Developmental Delay	<50	NR	NR	≥7,070	732.15	37.04
		Emotional Disturbance	<50	NR	NR	≥7,060	732.21	37.04
		HI—Deaf	<50	NR	NR	≥7,070	732.16	37.04
		HI—Hard-of-Hearing	<50	NR	NR	≥7,070	732.18	37.03
		Mild Mental Disability	<50	NR	NR	≥7,030	732.40	36.97
		Moderate Mental Disability	<50	NR	NR	≥7,070	732.15	37.04
		Orthopedic Impairment	<50	NR	NR	≥7,070	732.18	37.04
		Other Health Impairment	≥110	698.94	26.57	≥6,960	732.67	36.95
		Specific Learning Disability	≥340	696.61	20.66	≥6,730	733.97	36.78
		Speech or Language Impairment	<50	NR	NR	≥7,050	732.22	37.03
		Traumatic Brain Injury	<50	NR	NR	≥7,070	732.15	37.04
		Visual Impairment	<50	NR	NR	≥7,070	732.16	37.03
		Other	<50	NR	NR	≥7,070	732.17	37.04
Spring 2021	E	Gifted	≥1,160	794.24	24.16	≥45,560	737.97	35.54
		Talented	≥1,480	769.04	29.94	≥45,240	738.40	36.16
		Autism	≥240	719.41	35.50	≥46,490	739.48	36.35
		Deaf-Blindness	<50	NR	NR	≥46,730	739.38	36.38
		Developmental Delay	<50	NR	NR	≥46,720	739.38	36.37
		Emotional Disturbance	≥200	705.80	30.83	≥46,520	739.53	36.33
		HI—Deaf	<50	NR	NR	≥46,710	739.39	36.37
		HI—Hard-of-Hearing	≥50	715.61	31.81	≥46,670	739.41	36.37
		Mild Mental Disability	≥190	685.31	19.75	≥46,540	739.60	36.26
		Moderate Mental Disability	<50	NR	NR	≥46,720	739.39	36.37
		Orthopedic Impairment	<50	NR	NR	≥46,680	739.39	36.38
		Other Health Impairment	≥1,030	708.60	30.60	≥45,700	740.07	36.19

Admin.	Form	Group	Yes			No		
			N	Mean	Std. Dev.	N	Mean	Std. Dev.
		Specific Learning Disability	≥2,710	701.02	25.38	≥44,020	741.74	35.62
		Speech or Language Impairment	≥170	729.83	38.51	≥46,550	739.41	36.36
		Traumatic Brain Injury	<50	NR	NR	≥46,710	739.39	36.37
		Visual Impairment	<50	NR	NR	≥46,700	739.40	36.37
		Other	<50	NR	NR	≥46,720	739.38	36.37
Summer 2021	A	Gifted	<50	NR	NR	≥2,580	695.94	24.04
		Talented	<50	NR	NR	≥2,560	695.86	23.89
		Autism	<50	NR	NR	≥2,560	695.92	24.06
		Deaf-Blindness	<50	NR	NR	≥2,580	695.97	24.07
		Developmental Delay	<50	NR	NR	≥2,580	695.97	24.07
		Emotional Disturbance	<50	NR	NR	≥2,550	696.05	24.12
		HI—Deaf	<50	NR	NR	≥2,580	695.97	24.07
		HI—Hard-of-Hearing	<50	NR	NR	≥2,570	695.98	24.07
		Mild Mental Disability	<50	NR	NR	≥2,530	696.24	24.1
		Moderate Mental Disability	<50	NR	NR	≥2,570	696.00	24.06
		Orthopedic Impairment	<50	NR	NR	≥2,570	695.99	24.07
		Other Health Impairment	≥110	687.01	21.04	≥2,460	696.39	24.12
		Specific Learning Disability	≥390	687.83	18.38	≥2,180	697.45	24.68
		Speech or Language Impairment	<50	NR	NR	≥2,560	696.05	24.09
		Traumatic Brain Injury	<50	NR	NR	≥2,580	695.95	24.06
		Visual Impairment	<50	NR	NR	≥2,580	695.98	24.07
		Other	<50	NR	NR	≥2,580	695.97	24.07

Table 10.19 Special Education Classification Scale Score Means and Standard Deviations: English II

Special Education Classification Scale Score Means and Standard Deviations: English II								
Admin.	Form	Group	Yes			No		
			N	Mean	Std. Dev.	N	Mean	Std. Dev.
Fall 2020	B	Gifted	≥160	804.93	33.37	≥8,710	735.40	41.34
		Talented	≥180	768.31	34.17	≥8,700	736.04	42.17
		Autism	<50	NR	NR	≥8,840	736.82	42.21
		Deaf-Blindness	<50	NR	NR	≥8,880	736.70	42.27
		Developmental Delay	<50	NR	NR	≥8,870	736.72	42.26
		Emotional Disturbance	<50	NR	NR	≥8,860	736.77	42.25
		HI—Deaf	<50	NR	NR	≥8,880	736.70	42.27
		HI—Hard-of-Hearing	<50	NR	NR	≥8,870	736.73	42.25
		Mild Mental Disability	<50	NR	NR	≥8,850	736.89	42.18
		Moderate Mental Disability	<50	NR	NR	≥8,870	736.71	42.26
		Orthopedic Impairment	<50	NR	NR	≥8,870	736.71	42.26
		Other Health Impairment	≥80	700.72	32.78	≥8,790	737.04	42.20
		Specific Learning Disability	≥220	688.16	24.59	≥8,650	737.96	41.89
		Speech or Language Impairment	<50	NR	NR	≥8,860	736.72	42.27
		Traumatic Brain Injury	<50	NR	NR	≥8,880	736.70	42.27
		Visual Impairment	<50	NR	NR	≥8,870	736.69	42.28
		Other	<50	NR	NR	≥8,870	736.70	42.27
Spring 2021	D	Gifted	≥1,090	813.64	30.87	≥39,730	742.00	45.02
		Talented	≥1,240	776.91	38.55	≥39,590	742.89	46.02
		Autism	≥230	716.01	47.53	≥40,600	744.09	46.12
		Deaf-Blindness	<50	NR	NR	≥40,830	743.93	46.18
		Developmental Delay	<50	NR	NR	≥40,830	743.93	46.18
		Emotional Disturbance	≥130	700.97	39.43	≥40,700	744.07	46.14
		HI—Deaf	<50	NR	NR	≥40,810	743.94	46.17
		HI—Hard-of-Hearing	≥50	718.00	46.02	≥40,780	743.96	46.17
		Mild Mental Disability	≥140	671.63	19.83	≥40,690	744.18	46.05
		Moderate Mental Disability	<50	NR	NR	≥40,830	743.93	46.18
		Orthopedic Impairment	<50	NR	NR	≥40,790	743.93	46.18
		Other Health Impairment	≥820	700.58	36.80	≥40,010	744.82	45.93
		Specific Learning Disability	≥2,090	695.65	31.73	≥38,740	746.53	45.40
		Speech or Language Impairment	≥110	716.22	40.42	≥40,720	744.00	46.17

Special Education Classification Scale Score Means and Standard Deviations: English II								
Admin.	Form	Group	Yes			No		
			N	Mean	Std. Dev.	N	Mean	Std. Dev.
		Traumatic Brain Injury	<50	NR	NR	≥40,830	743.93	46.18
		Visual Impairment	<50	NR	NR	≥40,810	743.93	46.18
		Other	<50	NR	NR	≥40,830	743.93	46.18
Summer 2021	B	Gifted	<50	NR	NR	≥2,630	691.36	28.56
		Talented	<50	NR	NR	≥2,610	691.27	28.55
		Autism	<50	NR	NR	≥2,610	691.57	28.70
		Deaf-Blindness	<50	NR	NR	≥2,640	691.45	28.68
		Developmental Delay	<50	NR	NR	≥2,640	691.45	28.68
		Emotional Disturbance	<50	NR	NR	≥2,620	691.58	28.70
		HI—Deaf	<50	NR	NR	≥2,640	691.44	28.68
		HI—Hard-of-Hearing	<50	NR	NR	≥2,640	691.42	28.66
		Mild Mental Disability	<50	NR	NR	≥2,610	691.66	28.68
		Moderate Mental Disability	<50	NR	NR	≥2,640	691.47	28.68
		Orthopedic Impairment	<50	NR	NR	≥2,630	691.42	28.67
		Other Health Impairment	≥120	678.85	21.25	≥2,510	692.08	28.86
		Specific Learning Disability	≥330	678.07	22.39	≥2,300	693.40	28.98
		Speech or Language Impairment	<50	NR	NR	≥2,620	691.50	28.71
		Traumatic Brain Injury	<50	NR	NR	≥2,640	691.44	28.68
		Visual Impairment	<50	NR	NR	≥2,640	691.44	28.69
		Other	<50	NR	NR	≥2,640	691.45	28.68

Table 10.20 Special Education Classification Scale Score Means and Standard Deviations: Algebra I

Special Education Classification Scale Score Means and Standard Deviations: Algebra I								
Admin.	Form	Group	Yes			No		
			N	Mean	Std. Dev.	N	Mean	Std. Dev.
Fall 2020	D	Gifted	<50	NR	NR	≥5,070	727.99	30.74
		Talented	≥50	745.49	31.20	≥5,050	728.06	30.92
		Autism	<50	NR	NR	≥5,090	728.28	30.97
		Deaf-Blindness	<50	NR	NR	≥5,100	728.25	30.97
		Developmental Delay	<50	NR	NR	≥5,100	728.25	30.98
		Emotional Disturbance	<50	NR	NR	≥5,080	728.36	30.96
		HI—Deaf	<50	NR	NR	≥5,100	728.25	30.97
		HI—Hard-of-Hearing	<50	NR	NR	≥5,100	728.29	30.96
		Mild Mental Disability	<50	NR	NR	≥5,080	728.37	30.93

Special Education Classification Scale Score Means and Standard Deviations: Algebra I								
Admin.	Form	Group	Yes			No		
			N	Mean	Std. Dev.	N	Mean	Std. Dev.
		Moderate Mental Disability	<50	NR	NR	≥5,100	728.26	30.97
		Orthopedic Impairment	<50	NR	NR	≥5,100	728.27	30.97
		Other Health Impairment	≥80	704.50	21.46	≥5,020	728.65	30.95
		Specific Learning Disability	≥170	702.35	19.12	≥4,930	729.16	30.92
		Speech or Language Impairment	<50	NR	NR	≥5,090	728.30	30.96
		Traumatic Brain Injury	<50	NR	NR	≥5,100	728.25	30.97
		Visual Impairment	<50	NR	NR	≥5,100	728.25	30.97
		Other	<50	NR	NR	≥5,100	728.25	30.98
Spring 2021	E	Gifted	≥1,300	791.35	30.63	≥47,190	731.73	33.08
		Talented	≥1,600	756.79	32.59	≥46,890	732.53	34.17
		Autism	≥250	718.68	35.48	≥48,240	733.41	34.38
		Deaf-Blindness	<50	NR	NR	≥48,500	733.34	34.40
		Developmental Delay	<50	NR	NR	≥48,500	733.34	34.40
		Emotional Disturbance	≥220	705.74	27.24	≥48,280	733.46	34.38
		HI—Deaf	<50	NR	NR	≥48,480	733.34	34.40
		HI—Hard-of-Hearing	≥60	722.05	29.19	≥48,440	733.35	34.40
		Mild Mental Disability	≥190	696.85	16.81	≥48,300	733.49	34.37
		Moderate Mental Disability	<50	NR	NR	≥48,500	733.34	34.40
		Orthopedic Impairment	≥50	717.58	34.20	≥48,450	733.35	34.39
		Other Health Impairment	≥1,040	706.58	26.29	≥47,450	733.93	34.32
		Specific Learning Disability	≥2,710	703.82	21.11	≥45,780	735.09	34.24
		Speech or Language Impairment	≥190	730.73	36.40	≥48,310	733.35	34.39
		Traumatic Brain Injury	<50	NR	NR	≥48,490	733.34	34.40
		Visual Impairment	<50	NR	NR	≥48,470	733.34	34.39
		Other	<50	NR	NR	≥48,500	733.34	34.40
Summer 2021	D	Gifted	<50	NR	NR	≥3,520	707.95	21.93
		Talented	<50	NR	NR	≥3,500	708.03	22.31
		Autism	<50	NR	NR	≥3,510	708.24	22.48
		Deaf-Blindness	<50	NR	NR	≥3,540	708.23	22.52
		Developmental Delay	<50	NR	NR	≥3,540	708.23	22.52
		Emotional Disturbance	<50	NR	NR	≥3,500	708.33	22.54
		HI—Deaf	<50	NR	NR	≥3,540	708.22	22.52

Special Education Classification Scale Score Means and Standard Deviations: Algebra I								
Admin.	Form	Group	Yes			No		
			N	Mean	Std. Dev.	N	Mean	Std. Dev.
		HI—Hard-of-Hearing	<50	NR	NR	≥3,530	708.24	22.52
		Mild Mental Disability	<50	NR	NR	≥3,500	708.38	22.50
		Moderate Mental Disability	<50	NR	NR	≥3,530	708.23	22.52
		Orthopedic Impairment	<50	NR	NR	≥3,530	708.25	22.53
		Other Health Impairment	≥140	698.70	17.91	≥3,390	708.63	22.60
		Specific Learning Disability	≥380	698.41	17.72	≥3,150	709.43	22.75
		Speech or Language Impairment	<50	NR	NR	≥3,520	708.22	22.54
		Traumatic Brain Injury	<50	NR	NR	≥3,540	708.23	22.52
		Visual Impairment	<50	NR	NR	≥3,540	708.22	22.51
		Other	<50	NR	NR	≥3,540	708.23	22.52

Table 10.21 Special Education Classification Scale Score Means and Standard Deviations: Geometry

Special Education Classification Scale Score Means and Standard Deviations: Geometry								
Admin.	Form	Group	Yes			No		
			N	Mean	Std. Dev.	N	Mean	Std. Dev.
Fall 2020	D	Gifted	≥120	772.62	25.18	≥5,560	732.51	26.08
		Talented	≥130	749.80	26.58	≥5,540	732.97	26.58
		Autism	<50	NR	NR	≥5,670	733.36	26.69
		Deaf-Blindness	<50	NR	NR	≥5,680	733.37	26.70
		Developmental Delay	<50	NR	NR	≥5,680	733.37	26.70
		Emotional Disturbance	<50	NR	NR	≥5,670	733.43	26.70
		HI—Deaf	<50	NR	NR	≥5,680	733.38	26.71
		HI—Hard-of-Hearing	<50	NR	NR	≥5,680	733.39	26.70
		Mild Mental Disability	<50	NR	NR	≥5,680	733.39	26.69
		Moderate Mental Disability	<50	NR	NR	≥5,680	733.39	26.70
		Orthopedic Impairment	<50	NR	NR	≥5,680	733.38	26.70
		Other Health Impairment	<50	NR	NR	≥5,650	733.46	26.71
		Specific Learning Disability	≥60	711.19	14.63	≥5,620	733.62	26.70
		Speech or Language Impairment	<50	NR	NR	≥5,680	733.39	26.70
		Traumatic Brain Injury	<50	NR	NR	≥5,680	733.38	26.69

Special Education Classification Scale Score Means and Standard Deviations: Geometry								
Admin.	Form	Group	Yes			No		
			N	Mean	Std. Dev.	N	Mean	Std. Dev.
		Visual Impairment	<50	NR	NR	≥5,680	733.37	26.70
		Other	<50	NR	NR	≥5,680	733.37	26.70
Spring 2021	E	Gifted	≥1,100	775.09	24.98	≥33,150	732.84	25.55
		Talented	≥1,270	747.26	25.06	≥32,980	733.70	26.53
		Autism	≥130	724.21	27.50	≥34,130	734.25	26.59
		Deaf-Blindness	<50	NR	NR	≥34,260	734.21	26.60
		Developmental Delay	<50	NR	NR	≥34,260	734.21	26.60
		Emotional Disturbance	≥60	712.45	22.11	≥34,190	734.25	26.59
		HI—Deaf	<50	NR	NR	≥34,250	734.21	26.60
		HI—Hard-of-Hearing	<50	NR	NR	≥34,220	734.21	26.61
		Mild Mental Disability	≥50	701.78	18.78	≥34,210	734.26	26.58
		Moderate Mental Disability	<50	NR	NR	≥34,250	734.21	26.60
		Orthopedic Impairment	<50	NR	NR	≥34,230	734.21	26.60
		Other Health Impairment	≥430	713.73	21.18	≥33,820	734.47	26.56
		Specific Learning Disability	≥1,050	711.03	17.28	≥33,210	734.94	26.52
		Speech or Language Impairment	≥80	727.64	28.61	≥34,170	734.22	26.60
		Traumatic Brain Injury	<50	NR	NR	≥34,250	734.21	26.60
		Visual Impairment	<50	NR	NR	≥34,230	734.21	26.60
		Other	<50	NR	NR	≥34,260	734.21	26.60
Summer 2021	D	Gifted	<50	NR	NR	≥1,050	711.21	17.26
		Talented	<50	NR	NR	≥1,040	711.14	17.26
		Autism	<50	NR	NR	≥1,050	711.25	17.31
		Deaf-Blindness	<50	NR	NR	≥1,050	711.27	17.30
		Developmental Delay	<50	NR	NR	≥1,050	711.27	17.30
		Emotional Disturbance	<50	NR	NR	≥1,050	711.27	17.33
		HI—Deaf	<50	NR	NR	≥1,050	711.27	17.30
		HI—Hard-of-Hearing	<50	NR	NR	≥1,050	711.27	17.30
		Mild Mental Disability	<50	NR	NR	≥1,050	711.35	17.26
		Moderate Mental Disability	<50	NR	NR	≥1,050	711.28	17.30
		Orthopedic Impairment	<50	NR	NR	≥1,050	711.32	17.28
		Other Health Impairment	<50	NR	NR	≥1,020	711.46	17.23
		Specific Learning Disability	≥70	703.28	18.13	≥980	711.88	17.09

Special Education Classification Scale Score Means and Standard Deviations: Geometry								
Admin.	Form	Group	Yes			No		
			N	Mean	Std. Dev.	N	Mean	Std. Dev.
		Speech or Language Impairment	<50	NR	NR	≥1,050	711.25	17.34
		Traumatic Brain Injury	<50	NR	NR	≥1,050	711.27	17.30
		Visual Impairment	<50	NR	NR	≥1,050	711.28	17.30
		Other	<50	NR	NR	≥1,050	711.31	17.29

10.6 Summary

In summary, the overall purpose of this chapter is to address fairness concerns that are relevant to the administration of LEAP 2025 assessments. The information in this chapter addresses multiple best practices of the testing industry and is particularly related to the following standards:

Standard 3.1 Those responsible for test development, revision, and administration should design all steps of the testing process to promote valid score interpretations for intended score uses for the widest possible range of individuals and relevant subgroups in the intended population (63).

Standard 3.2 Test developers are responsible for developing tests that measure the intended construct and for minimizing the potential for tests' being affected by construct-irrelevant characteristics, such as linguistic, communicative, cognitive, cultural, physical, or other characteristics (64).

Standard 3.3 Those responsible for test development should include relevant subgroups in validity, reliability/precision, and other preliminary studies used when constructing the test (64).

Standard 3.4 Test takers should receive comparable treatment during the test administration and scoring process (65).

Standard 3.5 Test developers should specify and document provisions that have been made to test administration and scoring procedures to remove construct-irrelevant barriers for all relevant subgroups in the test-taker population (65).

Standard 3.6 Where credible evidence indicates that test scores may differ in meaning for relevant subgroups in the intended examinee population, test developers and/or users are responsible for examining the evidence for validity of score interpretations for intended uses for individuals from those subgroups. What constitutes a significant difference in subgroup scores and what actions are taken in response to such differences may be defined by applicable laws (65).

Standard 3.16 When credible research indicates that test scores for some relevant subgroups are differentially affected by construct-irrelevant characteristics of the test or of the examinees, when legally permissible, test users should use the test only for those subgroups for which there is sufficient evidence of validity to support score interpretations for the intended uses (70).

Appendix A—Accommodated Print Form Creation

Guidelines for Building Accommodated Print Forms

Careful consideration is given to all items that are used for accommodated print (AP) forms and/or braille forms. Fairness for all populations, item integrity, and student-item interaction for technology-enhanced (TE) items are all factors when selecting the items that will appear on an AP form. TE items are modified so that students who interact with an item on an AP form will have a similar experience to students who interact with that same item in the online environment. This maintains both the rigor and the content being assessed. Some examples of the modification process are provided below.

- Drag-and-drop items in the online environment require a student to place the answer options in an interactive table. For the AP form, the student is presented with a table with the same information as the interactive table (column or row headers, any completed cells, and blank spaces) and the answer options are listed below the table (similar to the online form in which the options are listed either below or to the right of the table). The directions are modified to ask the student to write the correct answer in its corresponding box. Students are also able to circle the text and draw arrows to indicate where it should be placed or add labels to the answer choices and write only the label in the box, as long as the intended response is clear to the test administrator who will transcribe the answers into the online system.
- Matching items in the online environment require a student to select a checkbox in one or more columns for each of multiple rows. In the AP form, the student is provided with a table and asked to mark an *X* in the correct places.
- Highlight-text items or item parts in the online environment require a student to click on the selected text, which highlights the selected word, phrase, or sentence. In the AP form, the text is presented in the same format and the student is asked to circle the answer. Where only certain words or phrases are selectable in the online system, those options are underlined in the AP form to indicate which words and/or phrases the student should select from.
- Drop-down menu items in the online environment have answer options in a drop-down menu format, oftentimes as part of a complete sentence. The AP form displays the item with a blank line in place of the drop-down menu in the sentence, with all the answer options for the drop-down menu presented vertically below the sentence. The directions are then modified to ask the student to circle the word/phrase that belongs in the blank.
- Short answer items in the online environment require a student to type the answer in a box. In the AP form, a box is provided for the student to write the response.
- Keypad input items in the online environment require a student to enter a numeric response including all rational and irrational numbers as well as expressions and equations. In the AP form, a box is provided for the student to write the response.
- Graphing items, including coordinate planes, number lines, line plots, and bar graphs, in the online environment require a student to complete a graph by plotting points, adding *X*s to create a line plot, or raising/lowering bars to create a bar graph or histogram. In the AP form, the student is provided with the same coordinate plane, number line, line plot, or bar graph as in the online item, including titles, axis labels, and keys, and is asked to complete the graph.

Displaying items similarly in both accommodated print forms and the online environment (and allowing students to interact with the items in a similar manner) maintains item integrity by assessing a similar

construct in a similar manner regardless of where a student encounters an item. This provides students who are unable to access the assessment online with an assessment at the same level of rigor as the online test.

AP forms are thoroughly reviewed by DRC and LDOE content experts to ensure a valid and reliable assessment for students who are unable to participate in the online assessment. These forms are also used as the source files for the creation of braille forms.

Appendix B—Transadaptation Process for Spanish Mathematics Forms

For English Learners, the LDOE offers the mathematics assessments in Spanish for computer-based tests (CBTs) in order to mirror the English language forms and the text-to-speech (TTS) forms. The Spanish-language versions of the test were developed through transadaptation. Transadaptation takes into consideration the grade-level appropriateness of the words and sentence structures used and the linguistic and cultural differences that exist between speakers of two different languages. Accounting for these differences allows experts to ensure that a Spanish-language version of an item will measure the same construct as the English-language version of the item at the same level of rigor. The item is, therefore, expected to measure the achievement of English Learners in the same way that the English version of the item does for native speakers of English.

Once the operational form was approved in English, DRC provided item IDs for acquired items to New Meridian, who then identified which of those items had previously appeared on a Spanish transadapted form. Once New Meridian identified the items that had previously been transadapted and provided the transadaptations of those items, DRC identified the English version of all items that had not been previously transadapted (either because they were Louisiana-owned items that would appear in field-test positions or because they were acquired items that had not been previously used on a New Meridian Spanish-language form). These items were then provided to the Spanish transadaptation subcontractor for initial transadaptation. DRC's Spanish Test Development team (who are all native Spanish speakers) reviewed the previously transadapted items to ensure consistency between those items transadapted as part of the New Meridian assessments and those transadapted specifically for Louisiana. The team provided guidance to the translator conducting the initial transadaptation in grade-level and culturally appropriate ways. Upon completion of the transadaptation by the subcontractor, DRC's Spanish Test Development team conducted reviews by native Spanish speakers for content and grade-level appropriateness of the transadaptation. The team also conducted an editorial review. At least two members of DRC's Spanish Test Development team compared each English item to the Spanish transadaptation to ensure that the transadaptation

- was accurate;
- contained grade-appropriate wording;
- contained answer choices that were reasonably parallel;
- did not introduce ambiguity into the Spanish version;
- contained graphics that were clearly transadapted;
- did not alter current teaching and learning practices in the content area; and
- remained free of gender, ethnic, cultural, socioeconomic, and regional bias.

The Spanish Test Development team then reconciled any discrepancies and submitted the transadaptations to a senior Spanish Test Development team member for resolution. After approval by the senior Spanish Test Development team member, the item moved forward to be imported into DRC's item banking system.

Both previously transadapted items and newly transadapted items were imported into DRC's item banking system and formatted for online use. Each Spanish item was paired with the corresponding English item in the item bank, and the Spanish item was formatted. Graphics for the item were then finalized for review. The

finalized transadaptation was then compared to the Spanish version of the item in the DRC assessment system and the English version of the item, and all changes were verified.

DRC's Spanish Test Development team then used the final, approved communication assistance scripts in English to transadapt descriptions of graphics as necessary. These descriptions were used when preparing the TTS forms for review. Scripting the TTS forms and reviewing the finalized Spanish forms were conducted by native Spanish speakers at DRC prior to submitting the forms to the LDOE for a translation review by a third-party translation vendor. The vendor reviewed the transadapted forms and provided feedback to the LDOE and DRC. Experienced DRC Spanish Test Development team members and the translation vendor resolved any issues, and DRC made modifications as necessary. The forms were then approved by both DRC and the LDOE translation vendor.

Appendix C—LEAP 2025 Spring 2021 Handscoring/AI Documentation

Appendix D—Classical Item Statistics

Table D.1 Operational Item Statistics—English I Fall 2020 Administration

Item	Item Type	Total N	Adj. N	*p*-Value	Pbis	% Omit	% at 0	% at 1	% at 2	% at 3	% at 4
1	ESR	≥5,740	≥5,740	0.51	0.27	0.02	41.54	14.84	43.60		
2	ESR	≥5,740	≥5,730	0.40	0.42	0.09	56.37	7.42	36.13		
3	ESR	≥5,740	≥5,730	0.53	0.33	0.05	29.94	33.46	36.54		
4	TE	≥5,740	≥5,730	0.31	0.45	0.09	67.78	1.85	30.29		
5	ESR	≥5,740	≥5,730	0.45	0.53	0.07	38.41	33.69	27.83		
6	ESR	≥5,740	≥5,730	0.53	0.48	0.19	44.61	4.32	50.88		
7	ESR	≥5,740	≥5,730	0.73	0.43	0.10	16.20	21.20	62.50		
8	TE	≥5,740	≥5,730	0.36	0.44	0.14	45.29	36.81	17.77		
9	CR	≥5,740	≥5,650	0.39	0.81	1.13	13.53	35.29	32.21	14.79	2.63
10	CR	≥5,740	≥5,650	0.39	0.81	1.13	13.53	35.29	32.21	14.79	2.63
11	CR	≥5,740	≥5,650	0.52	0.79	1.13	13.88	34.63	32.31	17.63	
12	ESR	≥5,740	≥5,730	0.44	0.51	0.05	45.17	21.74	33.04		
13	MS	≥5,740	≥5,730	0.47	0.57	0.12	39.23	26.62	34.04		
14	TE	≥5,740	≥5,730	0.28	0.28	0.10	55.03	33.50	11.37		
15	MS	≥5,740	≥5,730	0.49	0.36	0.10	36.32	29.80	33.77		
16	CR	≥5,740	≥5,380	0.32	0.79	4.30	27.42	25.52	29.21	8.83	2.82
17	CR	≥5,740	≥5,380	0.39	0.80	4.30	25.57	33.20	27.35	7.68	
18	ESR	≥5,740	≥5,720	0.37	0.26	0.23	46.35	32.22	21.20		
19	ESR	≥5,740	≥5,720	0.38	0.45	0.30	59.26	4.74	35.71		
20	ESR	≥5,740	≥5,720	0.49	0.36	0.26	43.63	14.84	41.26		
21	ESR	≥5,740	≥5,720	0.53	0.43	0.31	40.83	11.06	47.80		
22	TE	≥5,740	≥5,680	0.32	0.48	0.98	45.15	44.99	8.88		
23	TE	≥5,740	≥5,680	0.59	0.55	0.96	19.68	41.06	38.30		
24	ESR	≥5,740	≥5,740	0.27	0.39	0.02	68.23	9.44	22.31		
25	TE	≥5,740	≥5,710	0.35	0.33	0.40	49.17	31.20	19.23		
26	MS	≥5,740	≥5,730	0.42	0.36	0.16	23.71	69.01	7.12		
27	ESR	≥5,740	≥5,730	0.42	0.31	0.14	47.59	20.10	32.17		
28	ESR	≥5,740	≥5,730	0.09	0.16	0.16	87.49	6.90	5.45		
29	ESR	≥5,740	≥5,730	0.41	0.40	0.14	56.75	4.55	38.56		
30	ESR	≥5,740	≥5,730	0.51	0.48	0.16	42.73	12.07	45.04		
31	ESR	≥5,740	≥5,720	0.51	0.56	0.23	38.13	22.21	39.44		
32	ESR	≥5,740	≥5,730	0.27	0.24	0.19	62.65	21.09	16.06		
33	ESR	≥5,740	≥5,720	0.37	0.28	0.21	54.73	17.16	27.90		

Table D.2 Operational Item Statistics—English II Fall 2020 Administration

Item	Item Type	Total N	Adj. N	*p*-Value	Pbis	% Omit	% at 0	% at 1	% at 2	% at 3	% at 4
1	ESR	≥8,310	≥8,310	0.77	0.40	0.06	16.97	12.55	70.42		
2	ESR	≥8,310	≥8,290	0.30	0.31	0.22	52.90	33.44	13.44		
3	ESR	≥8,310	≥8,300	0.60	0.34	0.18	17.76	44.50	37.56		
4	TE	≥8,310	≥8,290	0.43	0.31	0.22	21.33	70.71	7.74		
5	MS	≥8,310	≥8,290	0.50	0.45	0.22	35.12	29.88	34.78		
6	ESR	≥8,310	≥8,300	0.81	0.47	0.18	13.60	10.89	75.33		
7	MS	≥8,310	≥8,290	0.40	0.47	0.26	32.58	55.50	11.65		
8	ESR	≥8,310	≥8,300	0.59	0.50	0.19	32.04	17.81	49.96		
9	CR	≥8,310	≥8,100	0.43	0.81	1.78	10.64	29.90	35.55	17.71	3.64
10	CR	≥8,310	≥8,100	0.43	0.81	1.78	10.64	29.90	35.55	17.71	3.64
11	CR	≥8,310	≥8,100	0.53	0.80	1.78	12.97	30.64	36.06	17.78	
12	ESR	≥8,310	≥8,310	0.40	0.30	0.08	48.96	22.27	28.69		
13	ESR	≥8,310	≥8,290	0.31	0.35	0.24	62.74	11.88	25.14		
14	MS	≥8,310	≥8,300	0.46	0.28	0.19	12.87	82.15	4.80		
15	MS	≥8,310	≥8,300	0.28	0.45	0.19	53.99	35.33	10.50		
16	CR	≥8,310	≥7,880	0.31	0.82	3.68	28.53	27.35	27.82	9.98	1.15
17	CR	≥8,310	≥7,880	0.40	0.81	3.68	28.92	27.97	27.43	10.53	
18	MS	≥8,310	≥8,290	0.26	0.33	0.32	57.36	32.31	10.00		
19	MS	≥8,310	≥8,290	0.27	0.42	0.22	68.46	9.65	21.67		
20	MS	≥8,310	≥8,290	0.25	0.27	0.25	63.70	23.15	12.90		
21	MS	≥8,310	≥8,290	0.30	0.27	0.29	51.29	37.56	10.86		
22	TE	≥8,310	≥8,270	0.62	0.46	0.49	15.09	45.44	38.98		
23	MS	≥8,310	≥8,280	0.61	0.61	0.44	22.98	30.85	45.73		
24	ESR	≥8,310	≥8,310	0.73	0.38	0.05	20.91	12.06	66.98		
25	ESR	≥8,310	≥8,300	0.46	0.40	0.20	36.01	36.50	27.28		
26	ESR	≥8,310	≥8,300	0.57	0.55	0.11	25.89	33.22	40.78		
27	ESR	≥8,310	≥8,300	0.60	0.37	0.10	27.50	25.86	46.54		
28	MS	≥8,310	≥8,300	0.46	0.46	0.18	45.57	17.52	36.73		
29	TE	≥8,310	≥8,280	0.40	0.40	0.36	46.12	27.99	25.53		
30	MS	≥8,310	≥8,290	0.38	0.55	0.32	49.61	24.44	25.62		
31	ESR	≥8,310	≥8,290	0.61	0.53	0.22	35.24	7.73	56.81		
32	MS	≥8,310	≥8,290	0.23	0.35	0.25	63.60	26.84	9.31		
33	MS	≥8,310	≥8,300	0.41	0.54	0.20	33.14	51.18	15.47		

Table D.3 Operational Item Statistics—Algebra I Fall 2020 Administration

Item	Item Type	Total N	Adj. N	*p*-Value	Pbis	% Omit	% at 0	% at 1/A	% at 2/B	% at 3/C	% at 4/D	% at 5	% at 6
1	MC	≥4,290	≥4,280	0.48	0.41	0.12		18.63	21.71	48.40	11.13		
2	MS	≥4,290	≥4,280	0.28	0.42	0.16	72.07	27.77					
3	SA	≥4,290	≥4,220	0.23	0.47	1.58	76.26	22.15					
4	MPSR	≥4,290	≥4,290	0.22	0.33	0.00	68.95	17.35	13.70				
5	MPSR	≥4,290	≥4,280	0.24	0.32	0.14	55.63	40.65	3.59				
6	MC	≥4,290	≥4,270	0.38	0.13	0.49		15.21	15.86	38.18	30.26		
7	MC	≥4,290	≥4,270	0.47	0.26	0.54		23.18	46.31	19.36	10.62		
8	MC	≥4,290	≥4,290	0.36	0.19	0.07		35.73	26.16	23.15	14.88		
9	MC	≥4,290	≥4,280	0.29	0.20	0.12		35.64	29.09	15.33	19.82		
10	SA	≥4,290	≥4,290	0.56	0.54	0.02	10.16	20.48	23.41	27.02	18.91		
11	MC	≥4,290	≥4,290	0.31	0.28	0.00		11.67	30.75	45.17	12.42		
12	CR	≥4,290	≥4,290	0.51	0.58	0.05	16.84	27.02	42.12	13.98			
13	CR	≥4,290	≥4,000	0.24	0.54	4.03	53.65	18.22	14.86	6.64			
14	MC	≥4,290	≥4,280	0.80	0.32	0.12		5.96	79.80	8.62	5.50		
15	MC	≥4,290	≥4,280	0.30	0.13	0.12		19.36	10.76	39.44	30.33		
16	TE	≥4,290	≥4,280	0.18	0.47	0.21	81.53	18.26					
17	TE	≥4,290	≥4,280	0.40	0.52	0.21	60.07	39.72					
18	MC	≥4,290	≥4,280	0.40	0.34	0.14		13.44	17.17	28.86	40.39		
19	MC	≥4,290	≥4,280	0.31	0.26	0.19		30.58	30.72	28.07	10.44		
20	MC	≥4,290	≥4,280	0.29	0.20	0.16		28.74	35.08	22.04	13.98		
21	MPSR	≥4,290	≥4,280	0.38	0.33	0.26	39.60	45.28	14.86				
22	MS	≥4,290	≥4,280	0.31	0.36	0.19	69.16	30.65					
23	SA	≥4,290	≥4,110	0.18	0.55	4.10	64.38	27.72	3.80				
24	MC	≥4,290	≥4,280	0.25	0.12	0.12		11.76	32.94	30.21	24.97		
25	CR	≥4,290	≥3,960	0.09	0.61	5.33	68.62	11.32	5.50	3.10	2.17	1.35	0.26
26	CR	≥4,290	≥4,250	0.10	0.45	0.89	73.05	17.63	4.26	2.70	1.47		
27	MC	≥4,290	≥4,280	0.41	0.21	0.19		27.14	17.87	40.58	14.23		
28	MC	≥4,290	≥4,280	0.25	0.26	0.19		25.32	32.77	33.94	7.78		
29	MPSR	≥4,290	≥4,290	0.30	0.45	0.07	51.15	37.29	11.48				
30	MC	≥4,290	≥4,280	0.57	0.30	0.28		12.16	56.39	23.85	7.31		
31	SA	≥4,290	≥4,280	0.20	0.47	0.26	67.23	25.97	6.55				
32	MC	≥4,290	≥4,280	0.30	0.14	0.28		35.50	12.81	21.27	30.14		
33	MPSR	≥4,290	≥4,270	0.39	0.22	0.35	36.92	48.24	14.49				
34	TE	≥4,290	≥4,270	0.42	0.24	0.35	58.16	41.49					
35	MS	≥4,290	≥4,280	0.27	0.52	0.26	72.61	27.14					
36	MC	≥4,290	≥4,270	0.46	0.17	0.37		11.25	24.78	18.12	45.47		
37	CR	≥4,290	≥3,900	0.12	0.52	6.36	68.67	16.40	1.91	3.89			
38	CR	≥4,290	≥3,570	0.10	0.56	9.76	66.22	9.41	2.89	1.82	2.96		
39	CR	≥4,290	≥3,700	0.14	0.49	9.20	64.22	10.92	8.97	2.24			

Table D.4 Operational Item Statistics—Geometry Fall 2020 Administration

Item	Item Type	Total N	Adj. N	*p*-Value	Pbis	% Omit	% at 0	% at 1/A	% at 2/B	% at 3/C	% at 4/D	% at 5	% at 6
1	MPSR	≥5,600	≥5,600	0.34	0.20	0.02	43.80	45.08	11.11				
2	MC	≥5,600	≥5,590	0.42	0.24	0.12		24.94	41.87	17.78	15.28		
3	MC	≥5,600	≥5,590	0.65	0.38	0.07		14.23	7.18	65.31	13.21		
4	TE	≥5,600	≥5,580	0.43	0.63	0.25	56.65	43.10					
5	SA	≥5,600	≥5,500	0.33	0.54	1.73	65.52	32.74					
6	SA	≥5,600	≥5,490	0.27	0.51	1.82	71.90	26.28					
7	MPSR	≥5,600	≥5,560	0.40	0.14	0.73	34.82	49.99	14.46				
8	TE	≥5,600	≥5,590	0.30	0.60	0.05	70.20	29.74					
9	MC	≥5,600	≥5,590	0.60	0.34	0.14		59.86	12.32	14.94	12.73		
10	SA	≥5,600	≥5,590	0.34	0.65	0.12	26.51	31.60	26.32	11.12	4.32		
11	SA	≥5,600	≥5,570	0.16	0.59	0.54	83.09	16.37					
12	CR	≥5,600	≥5,170	0.30	0.64	5.21	48.44	12.62	23.07	8.32			
13	CR	≥5,600	≥5,010	0.06	0.55	7.25	77.54	8.37	3.09	0.52			
14	MC	≥5,600	≥5,590	0.48	0.30	0.04		4.61	8.00	39.39	47.97		
15	MC	≥5,600	≥5,590	0.40	0.36	0.20		13.68	39.58	35.94	10.61		
16	SA	≥5,600	≥5,580	0.40	0.59	0.37	59.81	39.81					
17	MC	≥5,600	≥5,590	0.33	0.32	0.11		33.40	14.23	37.78	14.48		
18	MS	≥5,600	≥5,590	0.28	0.57	0.14	71.40	28.46					
19	TE	≥5,600	≥5,590	0.36	0.28	0.09	36.73	53.45	9.73				
20	MC	≥5,600	≥5,580	0.40	0.27	0.21		9.07	25.71	25.37	39.64		
21	MPSR	≥5,600	≥5,590	0.52	0.33	0.18	24.64	47.19	28.00				
22	CR	≥5,600	≥5,010	0.16	0.70	7.52	67.93	5.39	4.71	4.55	6.89		
23	MPSR	≥5,600	≥5,590	0.46	0.35	0.14	28.74	49.74	21.37				
24	MC	≥5,600	≥5,580	0.35	0.48	0.30		20.66	27.53	16.94	34.57		
25	CR	≥5,600	≥5,190	0.17	0.75	5.46	46.99	23.76	9.02	4.87	5.37	1.84	0.84
26	MC	≥5,600	≥5,590	0.42	0.40	0.16		20.32	28.80	41.83	8.89		
27	SA	≥5,600	≥5,540	0.30	0.41	0.98	68.93	30.08					
28	MC	≥5,600	≥5,590	0.41	0.14	0.14		38.49	40.51	14.00	6.86		
29	MC	≥5,600	≥5,590	0.26	0.26	0.11		30.58	26.91	26.12	16.28		
30	MS	≥5,600	≥5,590	0.19	0.47	0.16	81.15	18.69					
31	TE	≥5,600	≥5,590	0.27	0.23	0.18	53.79	38.49	7.53				
32	SA	≥5,600	≥5,550	0.40	0.61	0.79	59.36	39.85					
33	SA	≥5,600	≥5,510	0.31	0.64	1.46	67.97	30.57					
34	TE	≥5,600	≥5,590	0.44	0.53	0.20	55.51	44.30					
35	MPSR	≥5,600	≥5,590	0.48	0.60	0.16	38.12	27.21	34.51				
36	TE	≥5,600	≥5,560	0.18	0.48	0.70	81.52	17.78					
37	CR	≥5,600	≥4,940	0.06	0.54	8.37	76.22	7.59	2.37	1.05	1.00		
38	CR	≥5,600	≥4,890	0.09	0.59	9.00	76.40	3.46	3.04	4.45			
39	CR	≥5,600	≥5,050	0.09	0.60	7.32	75.50	5.87	7.34	1.46			

Table D.5 Operational Item Statistics—English I Spring 2021 Administration

Item	Item Type	Total N	Adj. N	*p*-Value	Pbis	% Omit	% at 0	% at 1	% at 2	% at 3	% at 4
1	ESR	≥45,840	≥45,830	0.46	0.34	0.02	51.09	6.76	42.12		
2	MS	≥45,840	≥45,760	0.40	0.54	0.17	45.14	29.68	25.01		
3	TE	≥45,840	≥45,700	0.40	0.35	0.32	40.05	38.80	20.83		
4	ESR	≥45,840	≥45,720	0.56	0.31	0.27	37.93	12.76	49.03		
5	ESR	≥45,840	≥45,730	0.47	0.48	0.23	49.21	6.65	43.91		
6	ESR	≥45,840	≥45,740	0.68	0.38	0.23	30.37	2.53	66.88		
7	MS	≥45,840	≥45,660	0.37	0.51	0.39	50.90	23.11	25.59		
8	TE	≥45,840	≥45,610	0.32	0.53	0.50	40.92	52.94	5.64		
9	CR	≥45,840	≥44,610	0.32	0.80	1.63	20.59	36.06	33.19	7.01	0.46
10	CR	≥45,840	≥44,610	0.32	0.80	1.63	20.59	36.06	33.19	7.01	0.46
11	CR	≥45,840	≥44,610	0.44	0.77	1.63	18.39	37.60	34.23	7.08	
12	ESR	≥45,840	≥45,810	0.61	0.45	0.06	31.60	15.23	53.11		
13	TE	≥45,840	≥45,780	0.54	0.59	0.12	24.20	43.29	32.39		
14	MS	≥45,840	≥45,810	0.35	0.45	0.08	55.32	19.32	25.29		
15	TE	≥45,840	≥45,740	0.34	0.47	0.21	47.73	36.87	15.18		
16	CR	≥45,840	≥44,160	0.27	0.78	2.25	33.87	31.73	22.12	6.86	1.77
17	CR	≥45,840	≥44,160	0.36	0.77	2.25	30.86	34.48	22.74	8.25	
18	ESR	≥45,840	≥45,740	0.58	0.42	0.22	34.26	16.20	49.32		
19	ESR	≥45,840	≥45,770	0.40	0.45	0.15	55.97	7.25	36.63		
20	TE	≥45,840	≥45,770	0.58	0.56	0.16	37.40	9.73	52.71		
21	ESR	≥45,840	≥45,740	0.46	0.44	0.21	43.77	21.22	34.80		
22	TE	≥45,840	≥45,680	0.36	0.43	0.35	44.16	38.60	16.88		
23	MS	≥45,840	≥45,730	0.47	0.59	0.24	31.41	43.64	24.71		
24	ESR	≥45,840	≥45,820	0.78	0.38	0.04	18.37	7.73	73.86		
25	TE	≥45,840	≥45,790	0.42	0.21	0.11	51.34	12.96	35.59		
26	ESR	≥45,840	≥45,770	0.43	0.37	0.15	47.77	18.65	33.43		
27	ESR	≥45,840	≥45,780	0.52	0.37	0.13	41.37	13.46	45.03		
28	MS	≥45,840	≥45,770	0.45	0.37	0.16	34.63	39.65	25.56		
29	ESR	≥45,840	≥45,780	0.54	0.45	0.14	38.91	13.99	46.95		
30	ESR	≥45,840	≥45,750	0.41	0.34	0.20	47.98	22.53	29.29		
31	TE	≥45,840	≥45,750	0.47	0.34	0.19	48.85	7.93	43.03		
32	ESR	≥45,840	≥45,760	0.39	0.44	0.17	50.60	20.43	28.79		
33	TE	≥45,840	≥45,690	0.41	0.49	0.33	39.28	39.43	20.96		

Table D.6 Operational Item Statistics—English II Spring 2021 Administration

Item	Item Type	Total N	Adj. N	*p*-Value	Pbis	% Omit	% at 0	% at 1	% at 2	% at 3	% at 4
1	ESR	≥39,950	≥39,940	0.48	0.25	0.03	46.02	12.47	41.49		
2	TE	≥39,950	≥39,900	0.53	0.46	0.13	45.64	3.45	50.79		
3	MS	≥39,950	≥39,890	0.49	0.46	0.16	38.62	24.95	36.27		
4	ESR	≥39,950	≥39,890	0.36	0.41	0.15	40.76	45.76	13.33		
5	ESR	≥39,950	≥39,890	0.29	0.27	0.14	53.58	34.24	12.04		
6	ESR	≥39,950	≥39,880	0.69	0.35	0.17	17.58	27.66	54.60		
7	MS	≥39,950	≥39,850	0.28	0.45	0.25	54.62	34.41	10.73		
8	TE	≥39,950	≥39,870	0.61	0.44	0.19	31.84	14.19	53.78		
9	CR	≥39,950	≥38,880	0.42	0.83	1.72	15.70	24.79	35.16	18.38	3.28
10	CR	≥39,950	≥38,880	0.42	0.83	1.72	15.70	24.79	35.16	18.38	3.28
11	CR	≥39,950	≥38,880	0.56	0.80	1.72	12.68	26.62	38.10	19.93	
12	ESR	≥39,950	≥39,920	0.54	0.49	0.08	37.24	17.56	45.11		
13	ESR	≥39,950	≥39,890	0.58	0.38	0.14	28.47	27.58	43.82		
14	MS	≥39,950	≥39,870	0.35	0.42	0.21	54.31	20.62	24.86		
15	ESR	≥39,950	≥39,900	0.52	0.36	0.12	41.95	11.51	46.42		
16	CR	≥39,950	≥38,250	0.42	0.78	2.73	13.44	26.44	37.95	14.44	3.49
17	CR	≥39,950	≥38,250	0.52	0.78	2.73	15.30	28.04	36.54	15.87	
18	ESR	≥39,950	≥39,860	0.48	0.40	0.21	47.14	8.82	43.83		
19	ESR	≥39,950	≥39,900	0.31	0.33	0.12	64.45	8.88	26.56		
20	ESR	≥39,950	≥39,880	0.42	0.30	0.18	44.68	27.09	28.05		
21	TE	≥39,950	≥39,880	0.28	0.34	0.18	50.10	43.92	5.80		
22	TE	≥39,950	≥39,780	0.66	0.64	0.44	23.89	20.28	55.39		
23	ESR	≥39,950	≥39,840	0.47	0.43	0.26	40.44	24.73	34.56		
24	MS	≥39,950	≥39,930	0.54	0.49	0.05	30.23	31.55	38.17		
25	ESR	≥39,950	≥39,880	0.64	0.41	0.18	32.86	5.83	61.14		
26	ESR	≥39,950	≥39,890	0.44	0.36	0.14	50.77	9.79	39.30		
27	TE	≥39,950	≥39,870	0.43	0.38	0.20	39.89	34.72	25.20		
28	MS	≥39,950	≥39,890	0.50	0.47	0.15	37.85	23.17	38.83		
29	MS	≥39,950	≥39,890	0.43	0.53	0.15	41.67	29.70	28.49		
30	MS	≥39,950	≥39,870	0.42	0.44	0.19	23.45	69.04	7.32		
31	TE	≥39,950	≥39,810	0.41	0.47	0.34	37.35	43.01	19.30		
32	MS	≥39,950	≥39,860	0.29	0.39	0.23	57.05	27.16	15.56		
33	TE	≥39,950	≥39,750	0.32	0.50	0.49	48.36	37.98	13.17		

Table D.7 Operational Item Statistics—Algebra I Spring 2021 Administration

Item	Item Type	Total N	Adj. N	*p*-Value	Pbis	% Omit	% at 0	% at 1/A	% at 2/B	% at 3/C	% at 4/D	% at 5	% at 6
1	MC	≥47,920	≥47,890	0.49	0.44	0.06		18.04	21.50	48.80	11.61		
2	MS	≥47,920	≥47,880	0.29	0.46	0.08	70.67	29.25					
3	MC	≥47,920	≥47,850	0.33	0.14	0.15		14.38	32.73	26.44	26.31		
4	MC	≥47,920	≥47,860	0.22	0.28	0.11		21.50	11.02	43.57	23.79		
5	MPSR	≥47,920	≥47,860	0.26	0.37	0.12	52.43	42.48	4.97				
6	MPSR	≥47,920	≥47,820	0.34	0.37	0.20	43.10	44.78	11.92				
7	MC	≥47,920	≥47,720	0.48	0.30	0.41		21.91	47.98	19.27	10.42		
8	MC	≥47,920	≥47,820	0.38	0.28	0.19		37.89	24.32	23.06	14.53		
9	MC	≥47,920	≥47,840	0.31	0.27	0.16		15.80	18.14	35.42	30.48		
10	MPSR	≥47,920	≥47,900	0.49	0.60	0.04	10.20	26.36	30.72	21.77	10.91		
11	MC	≥47,920	≥47,850	0.52	0.30	0.15		19.04	17.55	51.74	11.53		
12	CR	≥47,920	≥45,330	0.31	0.66	3.33	46.51	16.10	23.84	8.15			
13	CR	≥47,920	≥42,940	0.12	0.63	6.12	69.79	9.60	6.75	3.48			
14	MC	≥47,920	≥47,890	0.81	0.30	0.06		5.04	81.22	9.40	4.28		
15	SA	≥47,920	≥46,050	0.24	0.56	3.89	72.61	23.49					
16	TE	≥47,920	≥47,850	0.20	0.54	0.13	80.07	19.80					
17	SA	≥47,920	≥47,820	0.41	0.62	0.20	44.08	28.83	26.90				
18	MC	≥47,920	≥47,840	0.43	0.38	0.17		13.22	16.61	27.15	42.85		
19	MS	≥47,920	≥47,770	0.09	0.42	0.31	91.15	8.54					
20	MC	≥47,920	≥47,770	0.33	0.22	0.31		32.95	31.48	21.50	13.75		
21	MPSR	≥47,920	≥47,850	0.38	0.46	0.13	40.75	43.15	15.97				
22	MS	≥47,920	≥47,810	0.34	0.42	0.21	65.56	34.22					
23	SA	≥47,920	≥46,590	0.30	0.58	2.76	68.17	29.06					
24	MC	≥47,920	≥47,770	0.24	0.17	0.31		13.39	31.90	30.54	23.86		
25	CR	≥47,920	≥45,110	0.10	0.69	3.86	67.15	11.75	6.58	3.70	2.54	2.08	0.34
26	CR	≥47,920	≥43,050	0.22	0.49	7.35	28.42	47.13	11.02	2.08	1.19		
27	MC	≥47,920	≥47,830	0.42	0.26	0.19		26.57	18.92	41.76	12.56		
28	MC	≥47,920	≥47,790	0.28	0.42	0.26		27.80	23.13	22.47	26.35		
29	MC	≥47,920	≥47,800	0.40	0.16	0.25		40.27	32.82	16.36	10.29		
30	MC	≥47,920	≥47,820	0.58	0.31	0.20		11.43	57.84	23.61	6.92		
31	SA	≥47,920	≥47,820	0.22	0.53	0.20	65.63	25.37	8.79				
32	MC	≥47,920	≥47,840	0.41	0.18	0.16		7.63	24.91	26.74	40.56		
33	MPSR	≥47,920	≥47,870	0.42	0.36	0.10	30.10	54.90	14.89				
34	TE	≥47,920	≥47,470	0.31	0.54	0.93	51.32	33.22	14.53				
35	MC	≥47,920	≥47,810	0.42	0.16	0.22		12.31	41.78	19.13	26.56		
36	MC	≥47,920	≥47,810	0.31	0.41	0.22		23.05	19.73	26.47	30.54		
37	CR	≥47,920	≥44,400	0.24	0.62	4.85	53.32	18.40	14.39	6.56			
38	CR	≥47,920	≥41,910	0.11	0.63	8.05	67.78	10.34	3.45	1.54	4.36		
39	CR	≥47,920	≥43,330	0.11	0.61	6.53	70.25	12.34	4.72	3.12			

Table D.8 Operational Item Statistics—Geometry Spring 2021 Administration

Item	Item Type	Total N	Adj. N	*p*-Value	Pbis	% Omit	% at 0	% at 1/A	% at 2/B	% at 3/C	% at 4/D	% at 5	% at 6
1	TE	≥34,280	≥34,270	0.24	0.40	0.02	57.02	37.56	5.40				
2	MC	≥34,280	≥34,240	0.44	0.25	0.11		23.23	43.70	17.39	15.57		
3	MS	≥34,280	≥34,240	0.32	0.60	0.13	67.93	31.94					
4	TE	≥34,280	≥34,230	0.45	0.63	0.15	54.76	45.09					
5	TE	≥34,280	≥34,180	0.19	0.46	0.30	80.54	19.16					
6	SA	≥34,280	≥33,850	0.28	0.51	1.26	71.31	27.43					
7	MPSR	≥34,280	≥34,130	0.29	0.34	0.45	53.36	35.38	10.81				
8	MC	≥34,280	≥34,260	0.45	0.22	0.07		24.07	44.92	19.81	11.14		
9	MC	≥34,280	≥34,250	0.61	0.33	0.09		61.33	11.02	22.76	4.79		
10	MPSR	≥34,280	≥34,270	0.37	0.45	0.03	14.91	36.08	35.04	12.73	1.22		
11	SA	≥34,280	≥34,090	0.18	0.58	0.57	81.66	17.77					
12	CR	≥34,280	≥31,890	0.29	0.62	4.48	50.63	12.80	21.34	8.25			
13	CR	≥34,280	≥31,150	0.06	0.56	6.15	78.38	9.39	2.53	0.57			
14	MC	≥34,280	≥34,250	0.48	0.32	0.09		4.36	7.14	40.58	47.83		
15	MC	≥34,280	≥34,210	0.39	0.34	0.22		14.11	38.57	36.42	10.68		
16	SA	≥34,280	≥34,090	0.49	0.42	0.57	50.77	48.66					
17	MS	≥34,280	≥34,240	0.53	0.60	0.12	46.86	53.02					
18	MS	≥34,280	≥34,250	0.29	0.57	0.10	70.46	29.44					
19	TE	≥34,280	≥34,260	0.35	0.30	0.06	39.22	50.81	9.90				
20	MC	≥34,280	≥34,220	0.41	0.43	0.17		14.83	17.80	25.90	41.31		
21	MPSR	≥34,280	≥34,260	0.44	0.39	0.07	23.16	64.77	12.00				
22	CR	≥34,280	≥29,830	0.12	0.67	8.70	68.57	5.84	4.86	3.38	4.36		
23	MPSR	≥34,280	≥34,240	0.49	0.32	0.11	25.02	51.49	23.38				
24	MC	≥34,280	≥34,230	0.33	0.45	0.16		20.03	31.00	16.08	32.73		
25	CR	≥34,280	≥32,300	0.17	0.75	4.04	50.68	19.91	9.12	4.44	4.58	4.14	1.35
26	MC	≥34,280	≥34,240	0.51	0.51	0.13		15.42	29.73	51.09	3.63		
27	MC	≥34,280	≥34,210	0.41	0.32	0.20		9.50	30.69	18.59	41.03		
28	MC	≥34,280	≥34,220	0.48	0.25	0.18		24.65	47.42	14.06	13.69		
29	MC	≥34,280	≥34,230	0.24	0.31	0.16		24.42	29.84	22.95	22.63		
30	MC	≥34,280	≥34,220	0.31	0.31	0.17		23.18	25.04	20.89	30.72		
31	TE	≥34,280	≥34,250	0.26	0.24	0.10	54.36	38.17	7.37				
32	MC	≥34,280	≥34,230	0.37	0.30	0.16		16.61	36.53	21.42	25.29		
33	TE	≥34,280	≥34,240	0.28	0.52	0.13	72.25	27.62					
34	MC	≥34,280	≥34,240	0.47	0.28	0.13		13.60	47.12	18.41	20.75		
35	MPSR	≥34,280	≥34,250	0.49	0.60	0.09	37.56	26.78	35.57				
36	SA	≥34,280	≥33,750	0.12	0.60	1.56	87.01	11.44					
37	CR	≥34,280	≥34,200	0.14	0.73	0.23	68.58	14.98	7.86	7.13	1.22		
38	CR	≥34,280	≥29,990	0.09	0.57	8.27	77.05	3.12	2.55	4.76			
39	CR	≥34,280	≥31,200	0.13	0.61	6.40	65.57	17.39	6.36	1.69			

Table D.9 Operational Item Statistics—English I Summer 2021 Administration

Item	Item Type	Total N	Adj. N	p-Value	Pbis	% Omit	% at 0	% at 1	% at 2	% at 3	% at 4
1	ESR	≥2,610	≥2,610	0.21	0.23	0.08	68.07	21.98	9.88		
2	TE	≥2,610	≥2,590	0.09	0.38	0.73	85.49	10.15	3.64		
3	ESR	≥2,610	≥2,600	0.37	0.25	0.11	50.15	26.23	23.51		
4	ESR	≥2,610	≥2,600	0.18	0.28	0.31	72.93	18.57	8.19		
5	ESR	≥2,610	≥2,600	0.46	0.32	0.19	43.72	21.06	35.03		
6	ESR	≥2,610	≥2,600	0.29	0.26	0.27	62.75	16.73	20.25		
7	CR	≥2,610	≥2,350	0.05	0.65	5.74	73.51	15.12	1.34	0.08	0.04
8	CR	≥2,610	≥2,350	0.05	0.65	5.74	73.51	15.12	1.34	0.08	0.04
9	CR	≥2,610	≥2,350	0.09	0.59	5.74	66.69	21.59	1.68	0.11	
10	ESR	≥2,610	≥2,600	0.14	0.13	0.27	77.79	15.51	6.43		
11	TE	≥2,610	≥2,580	0.41	0.37	1.00	34.07	49.43	15.51		
12	ESR	≥2,610	≥2,600	0.35	0.40	0.42	54.02	22.01	23.55		
13	ESR	≥2,610	≥2,600	0.24	0.23	0.46	66.62	17.88	15.05		
14	ESR	≥2,610	≥2,610	0.30	0.18	0.08	60.30	19.60	20.02		
15	ESR	≥2,610	≥2,600	0.16	0.26	0.11	78.22	10.68	10.99		
16	ESR	≥2,610	≥2,610	0.36	0.18	0.08	47.66	32.62	19.64		
17	TE	≥2,610	≥2,600	0.10	0.22	0.23	86.75	6.39	6.62		
18	ESR	≥2,610	≥2,600	0.16	0.27	0.23	73.47	20.56	5.74		
19	ESR	≥2,610	≥2,600	0.20	0.30	0.31	75.61	8.46	15.62		
20	ESR	≥2,610	≥2,600	0.38	0.39	0.46	49.00	26.26	24.27		
21	TE	≥2,610	≥2,600	0.18	0.19	0.38	66.96	30.05	2.60		
22	CR	≥2,610	≥2,310	0.10	0.67	6.09	57.81	27.03	3.14	0.50	0.04
23	CR	≥2,610	≥2,310	0.10	0.67	6.09	57.81	27.03	3.14	0.50	0.04
24	CR	≥2,610	≥2,310	0.13	0.65	6.09	59.19	25.54	3.18	0.61	
25	MS	≥2,610	≥2,600	0.29	0.24	0.27	50.23	40.28	9.23		
26	ESR	≥2,610	≥2,600	0.31	0.33	0.46	59.53	17.96	22.05		
27	ESR	≥2,610	≥2,600	0.25	0.15	0.38	64.17	20.29	15.16		
28	TE	≥2,610	≥2,600	0.28	0.28	0.31	63.74	15.51	20.44		
29	ESR	≥2,610	≥2,600	0.20	0.26	0.31	69.83	19.03	10.83		
30	TE	≥2,610	≥2,590	0.16	0.35	0.80	71.55	22.82	4.82		
31	ESR	≥2,610	≥2,590	0.19	-0.08	0.57	62.67	34.99	1.76		
32	ESR	≥2,610	≥2,590	0.35	0.29	0.57	53.48	23.16	22.78		
33	ESR	≥2,610	≥2,590	0.16	0.21	0.54	74.46	18.15	6.85		
34	TE	≥2,610	≥2,580	0.15	0.29	1.23	73.70	21.02	4.06		

Table D.10 Operational Item Statistics—English II Summer 2021 Administration

Item	Item Type	Total N	Adj. N	p-Value	Pbis	% Omit	% at 0	% at 1	% at 2	% at 3	% at 4
1	ESR	≥2,650	≥2,650	0.23	0.24	0.08	70.52	13.63	15.78		
2	TE	≥2,650	≥2,650	0.25	0.26	0.15	59.94	29.78	10.13		
3	ESR	≥2,650	≥2,640	0.17	0.07	0.38	74.92	15.17	9.53		
4	MS	≥2,650	≥2,640	0.14	0.24	0.38	77.37	15.89	6.36		
5	MS	≥2,650	≥2,640	0.14	0.02	0.30	74.62	22.48	2.60		
6	TE	≥2,650	≥2,640	0.27	0.19	0.30	46.12	52.71	0.87		
7	CR	≥2,650	≥2,390	0.11	0.70	6.25	56.55	28.61	4.78	0.11	
8	CR	≥2,650	≥2,390	0.11	0.70	6.25	56.55	28.61	4.78	0.11	
9	CR	≥2,650	≥2,390	0.14	0.69	6.25	58.32	26.58	4.93	0.23	
10	ESR	≥2,650	≥2,640	0.24	0.20	0.49	66.23	19.09	14.19		
11	ESR	≥2,650	≥2,630	0.34	0.25	0.72	55.27	20.78	23.23		
12	MS	≥2,650	≥2,630	0.16	0.14	0.75	69.31	28.01	1.92		
13	TE	≥2,650	≥2,620	0.10	0.31	1.13	81.74	14.80	2.33		
14	ESR	≥2,650	≥2,650	0.53	0.33	0.11	39.57	14.19	46.12		
15	ESR	≥2,650	≥2,650	0.20	0.13	0.19	64.98	28.88	5.95		
16	ESR	≥2,650	≥2,640	0.39	0.29	0.34	42.66	35.43	21.57		
17	TE	≥2,650	≥2,640	0.29	0.28	0.26	43.26	55.38	1.09		
18	MS	≥2,650	≥2,650	0.25	0.32	0.23	57.87	33.09	8.81		
19	ESR	≥2,650	≥2,650	0.54	0.30	0.19	37.01	17.88	44.92		
20	MS	≥2,650	≥2,640	0.19	0.30	0.34	64.50	32.61	2.56		
21	ESR	≥2,650	≥2,650	0.27	0.31	0.23	63.18	20.07	16.53		
22	CR	≥2,650	≥2,390	0.14	0.72	5.95	47.67	36.45	5.46	0.49	0.04
23	CR	≥2,650	≥2,390	0.14	0.72	5.95	47.67	36.45	5.46	0.49	0.04
24	CR	≥2,650	≥2,390	0.16	0.69	5.95	52.60	32.38	4.71	0.41	
25	ESR	≥2,650	≥2,650	0.49	0.32	0.15	43.67	14.87	41.30		
26	ESR	≥2,650	≥2,640	0.24	0.26	0.45	60.77	29.07	9.71		
27	ESR	≥2,650	≥2,650	0.27	0.39	0.23	57.23	31.29	11.26		
28	ESR	≥2,650	≥2,640	0.33	0.28	0.26	55.76	22.48	21.50		
29	MS	≥2,650	≥2,640	0.27	0.18	0.38	56.33	32.76	10.54		
30	MS	≥2,650	≥2,640	0.18	0.27	0.26	70.75	21.61	7.38		
31	MS	≥2,650	≥2,640	0.26	0.16	0.34	48.72	49.66	1.28		
32	TE	≥2,650	≥2,640	0.23	0.19	0.38	57.64	38.55	3.43		
33	MS	≥2,650	≥2,640	0.18	0.04	0.34	67.51	29.29	2.86		
34	TE	≥2,650	≥2,640	0.15	0.21	0.45	70.07	28.28	1.20		

Table D.11 Operational Item Statistics—Algebra I Summer 2021 Administration

Item	Item Type	Total N	Adj. N	*p*-Value	Pbis	% Omit	% at 0	% at 1/A	% at 2/B	% at 3/C	% at 4/D	% at 5	% at 6
1	MC	≥3,550	≥3,550	0.28	0.20	0.17		24.44	35.12	27.87	12.40		
2	MS	≥3,550	≥3,550	0.13	0.23	0.14	86.78	13.08					
3	SA	≥3,550	≥3,500	0.05	0.47	1.49	93.64	4.87					
4	MPSR	≥3,550	≥3,550	0.15	0.08	0.06	73.26	22.78	3.91				
5	MPSR	≥3,550	≥3,540	0.19	0.21	0.20	64.65	33.01	2.14				
6	MC	≥3,550	≥3,540	0.37	0.04	0.45		12.71	15.52	36.75	34.56		
7	MC	≥3,550	≥3,520	0.36	0.19	0.79		25.84	35.57	27.14	10.66		
8	MC	≥3,550	≥3,540	0.30	0.15	0.22		29.78	29.89	21.15	18.95		
9	MC	≥3,550	≥3,540	0.26	0.09	0.22		26.97	26.18	24.83	21.79		
10	SA	≥3,550	≥3,550	0.37	0.36	0.11	20.53	33.94	27.14	13.95	4.33		
11	MC	≥3,550	≥3,550	0.26	0.11	0.14		16.54	25.96	44.91	12.46		
12	CR	≥3,550	≥3,550	0.32	0.35	0.14	33.80	39.34	23.28	3.43			
13	CR	≥3,550	≥3,150	0.09	0.46	6.38	71.57	12.23	3.74	1.07			
14	MC	≥3,550	≥3,540	0.61	0.23	0.28		12.15	60.97	16.79	9.81		
15	MC	≥3,550	≥3,540	0.28	0.07	0.31		20.98	16.51	33.83	28.37		
16	TE	≥3,550	≥3,540	0.07	0.39	0.28	93.14	6.58					
17	TE	≥3,550	≥3,530	0.14	0.38	0.48	85.26	14.26					
18	MC	≥3,550	≥3,540	0.24	0.24	0.42		19.46	18.36	37.88	23.88		
19	MC	≥3,550	≥3,530	0.19	0.16	0.48		19.15	32.48	34.06	13.84		
20	MC	≥3,550	≥3,530	0.22	0.10	0.48		22.38	32.68	27.08	17.38		
21	MPSR	≥3,550	≥3,540	0.30	0.19	0.34	46.99	45.36	7.31				
22	MS	≥3,550	≥3,540	0.21	0.23	0.39	78.68	20.92					
23	SA	≥3,550	≥3,370	0.06	0.49	4.98	84.11	9.65	1.27				
24	MC	≥3,550	≥3,540	0.20	0.05	0.45		11.90	38.53	29.47	19.66		
25	CR	≥3,550	≥3,160	0.03	0.62	6.81	81.75	4.36	1.21	0.59	0.65	0.28	0.14
26	CR	≥3,550	≥3,520	0.04	0.36	0.82	84.67	12.63	0.90	0.65	0.34		
27	MC	≥3,550	≥3,550	0.33	0.12	0.17		23.37	28.77	32.59	15.10		
28	MC	≥3,550	≥3,540	0.19	0.19	0.37		18.56	31.64	42.27	7.17		
29	MPSR	≥3,550	≥3,550	0.23	0.25	0.14	59.96	34.56	5.34				
30	MC	≥3,550	≥3,540	0.44	0.19	0.20		11.59	43.73	32.65	11.84		
31	SA	≥3,550	≥3,540	0.09	0.35	0.22	82.62	15.69	1.46				
32	MC	≥3,550	≥3,540	0.26	0.12	0.39		31.92	15.21	26.83	25.65		
33	MPSR	≥3,550	≥3,540	0.31	0.13	0.20	47.44	43.11	9.25				
34	TE	≥3,550	≥3,540	0.32	0.13	0.39	67.91	31.69					
35	MS	≥3,550	≥3,540	0.11	0.33	0.20	88.84	10.97					
36	MC	≥3,550	≥3,540	0.35	0.13	0.31		13.61	26.46	25.11	34.51		
37	CR	≥3,550	≥3,120	0.02	0.56	7.00	84.28	2.70	0.37	0.53			
38	CR	≥3,550	≥2,930	0.02	0.54	9.96	78.46	2.92	0.51	0.22	0.53		
39	CR	≥3,550	≥2,890	0.03	0.50	11.00	76.10	3.49	1.46	0.48			

Table D.12 Operational Item Statistics—Geometry Summer 2021 Administration

Item	Item Type	Total N	Adj. N	*p*-Value	Pbis	% Omit	% at 0	% at 1/A	% at 2/B	% at 3/C	% at 4/D	% at 5	% at 6
1	MPSR	≥1,060	≥1,060	0.29	0.14	0.00	50.38	40.60	9.02				
2	MC	≥1,060	≥1,060	0.35	0.12	0.38		34.59	34.40	21.05	9.59		
3	MC	≥1,060	≥1,060	0.45	0.19	0.19		21.62	12.31	44.83	21.05		
4	TE	≥1,060	≥1,060	0.11	0.49	0.28	89.19	10.53					
5	SA	≥1,060	≥1,040	0.09	0.39	1.88	89.57	8.55					
6	SA	≥1,060	≥1,040	0.10	0.33	1.60	88.63	9.77					
7	MPSR	≥1,060	≥1,050	0.36	0.05	0.75	41.45	45.02	12.78				
8	TE	≥1,060	≥1,060	0.06	0.44	0.09	94.27	5.64					
9	MC	≥1,060	≥1,060	0.35	0.22	0.00		35.34	18.42	22.18	24.06		
10	SA	≥1,060	≥1,060	0.18	0.36	0.19	46.99	36.47	13.44	1.79	1.13		
11	SA	≥1,060	≥1,050	0.02	0.54	0.47	97.74	1.79					
12	CR	≥1,060	≥940	0.10	0.37	6.58	72.37	8.08	7.42	0.85			
13	CR	≥1,060	≥930	0.01	0.47	7.14	86.56	0.66	0.28	0.28			
14	MC	≥1,060	≥1,060	0.30	0.19	0.00		9.30	13.63	47.18	29.89		
15	MC	≥1,060	≥1,050	0.29	0.11	0.47		26.22	28.57	31.86	12.88		
16	SA	≥1,060	≥1,050	0.14	0.25	0.47	85.34	14.19					
17	MC	≥1,060	≥1,060	0.21	0.15	0.19		21.05	21.62	38.44	18.70		
18	MS	≥1,060	≥1,060	0.08	0.36	0.19	91.64	8.18					
19	TE	≥1,060	≥1,060	0.29	0.07	0.00	49.62	43.42	6.95				
20	MC	≥1,060	≥1,060	0.32	0.11	0.00		12.50	20.21	35.24	32.05		
21	MPSR	≥1,060	≥1,060	0.40	0.14	0.00	37.31	44.92	17.76				
22	CR	≥1,060	≥930	0.01	0.52	7.33	85.81	0.94	0.28	0.28	0.28		
23	MPSR	≥1,060	≥1,060	0.38	0.15	0.00	35.15	53.85	11.00				
24	MC	≥1,060	≥1,060	0.16	0.19	0.00		26.88	33.83	23.40	15.88		
25	CR	≥1,060	≥960	0.03	0.40	6.11	76.32	12.22	1.69		0.38		
26	MC	≥1,060	≥1,060	0.27	0.12	0.00		20.30	41.54	27.07	11.09		
27	SA	≥1,060	≥1,040	0.19	0.20	1.50	79.51	18.98					
28	MC	≥1,060	≥1,060	0.33	0.02	0.09		33.08	33.27	22.84	10.71		
29	MC	≥1,060	≥1,060	0.17	0.02	0.09		38.25	22.93	17.29	21.43		
30	MS	≥1,060	≥1,060	0.06	0.07	0.19	93.89	5.92					
31	TE	≥1,060	≥1,060	0.22	0.00	0.09	59.87	36.37	3.67				
32	SA	≥1,060	≥1,050	0.07	0.42	1.22	91.54	7.24					
33	SA	≥1,060	≥1,040	0.06	0.36	1.41	92.48	6.11					
34	TE	≥1,060	≥1,060	0.17	0.25	0.09	82.61	17.29					
35	MPSR	≥1,060	≥1,060	0.20	0.31	0.09	67.01	25.94	6.95				
36	TE	≥1,060	≥1,050	0.07	0.23	0.56	92.39	7.05					
37	CR	≥1,060	≥890	0.01	0.37	9.21	80.83	2.63	0.09	0.09	0.09		
38	CR	≥1,060	≥920	0.01	0.54	7.71	86.09	0.47	0.28	0.28			

Appendix E—Student Participation Counts

Table E.1 Count of Students taking the Fall 2020 Administration: English I

Group	Grade							
	6	7	8	9	10	11	12	Total
All Students	<10	<10	≥10	≥4,950	≥850	≥760	≥480	≥7,070
Gender								
Female	<10	<10	<10	≥2,510	≥280	≥240	≥180	≥3,230
Male	<10	<10	<10	≥2,440	≥560	≥520	≥300	≥3,840
Ethnicity								
Hispanic/Latino	<10	<10	<10	≥400	≥90	≥140	≥130	≥780
American Indian or Alaska Native	<10	<10	<10	≥20	<10	<10	<10	≥30
Asian	<10	<10	<10	≥90	<10	≥10	<10	≥120
Black or African American	<10	<10	<10	≥2,200	≥460	≥450	≥280	≥3,410
Native Hawaiian or Other Pacific	<10	<10	<10	<10	<10	<10	<10	<10
White	<10	<10	<10	≥2,100	≥260	≥130	≥50	≥2,560
Two or More Races	<10	<10	<10	≥120	≥10	≥10	<10	≥150
Education Classification								
Regular	<10	<10	≥10	≥4,640	≥720	≥580	≥390	≥6,350
Special	<10	<10	<10	≥180	≥120	≥170	≥90	≥570
Gifted	<10	<10	<10	≥130	<10	<10	<10	≥140
Economic Status*								
Economically Disadvantaged	—	—	—	—	—	—	—	—
Not Economically Disadvantaged	—	—	—	—	—	—	—	—
English Learner Status								
Non-EL	<10	<10	≥10	≥4,820	≥750	≥610	≥340	≥6,540
EL	<10	<10	<10	≥130	≥100	≥150	≥140	≥530
Migrant Status								
Non-migrant	<10	<10	≥10	≥4,950	≥850	≥760	≥480	≥7,060
Migrant	<10	<10	<10	<10	<10	<10	<10	≥10
Section 504 Status								
Non-section 504	<10	<10	≥10	≥4,500	≥710	≥630	≥420	≥6,290
Section 504	<10	<10	<10	≥450	≥130	≥120	≥60	≥780
Homeless Status								
Not Homeless	<10	<10	≥10	≥4,920	≥840	≥740	≥470	≥7,000
Homeless	<10	<10	<10	≥30	≥10	≥10	≥10	≥70
Military Affiliation								
Not Military Affiliated	<10	<10	≥10	≥4,920	≥840	≥760	≥480	≥7,030
Military Affiliated	<10	<10	<10	≥30	<10	<10	<10	≥40
Foster Care Status								
Not in Foster Care	<10	<10	≥10	≥4,950	≥840	≥760	≥480	≥7,070
Foster Care	<10	<10	<10	<10	<10	<10	<10	<10

* Economic status information is not available for the fall and summer administrations.

Table E.2 Percentage of Students taking the Fall 2020 Administration: English I

Group	Grade							
	6	7	8	9	10	11	12	Total
All Students	0.00	0.00	0.20	70.05	12.04	10.82	6.89	100
Gender								
Female	0.00	0.00	0.25	77.78	8.88	7.49	5.60	100
Male	0.00	0.00	0.16	63.55	14.69	13.62	7.98	100
Ethnicity								
Hispanic/Latino	0.00	0.00	0.13	51.40	12.47	18.70	17.30	100
American Indian or Alaska Native	0.00	0.00	3.23	70.97	12.90	6.45	6.45	100
Asian	0.00	0.00	0.00	78.05	5.69	13.01	3.25	100
Black or African American	0.00	0.00	0.18	64.52	13.72	13.34	8.24	100
Native Hawaiian or Other Pacific	0.00	0.00	0.00	100.00	0.00	0.00	0.00	100
White	0.00	0.00	0.23	82.02	10.18	5.27	2.30	100
Two or More Races	0.00	0.00	0.00	80.50	8.81	6.92	3.77	100
Education Classification								
Regular	0.00	0.00	0.22	72.99	11.44	9.19	6.17	100
Special	0.00	0.00	0.00	31.65	21.22	30.78	16.35	100
Gifted	0.00	0.00	0.00	93.15	2.05	3.42	1.37	100
Economic Status*								
Economically Disadvantaged	—	—	—	—	—	—	—	—
Not Economically Disadvantaged	—	—	—	—	—	—	—	—
English Learner Status								
Non-EL	0.00	0.00	0.21	73.63	11.46	9.38	5.32	100
EL	0.00	0.00	0.00	25.94	19.17	28.57	26.32	100
Migrant Status								
Non-migrant	0.00	0.00	0.20	70.14	12.06	10.79	6.81	100
Migrant	0.00	0.00	0.00	26.67	0.00	26.67	46.67	100
Section 504 Status								
Non-section 504	0.00	0.00	0.22	71.60	11.36	10.15	6.67	100
Section 504	0.00	0.00	0.00	57.54	17.52	16.24	8.70	100
Homeless Status								
Not Homeless	0.00	0.00	0.19	70.34	11.99	10.68	6.80	100
Homeless	0.00	0.00	1.33	42.67	16.00	24.00	16.00	100
Military Affiliation								
Not Military Affiliated	0.00	0.00	0.17	70.06	12.05	10.84	6.88	100
Military Affiliated	0.00	0.00	4.26	68.09	10.64	8.51	8.51	100
Foster Care Status								
Not in Foster Care	0.00	0.00	0.20	70.10	12.01	10.79	6.90	100
Foster Care	0.00	0.00	0.00	14.29	42.86	42.86	0.00	100

* Economic status information is not available for the fall and summer administrations.

Table E.3 Count of Students taking the Fall 2020 Administration: English II

Group	Grade							
	6	7	8	9	10	11	12	Total
All Students	<10	<10	<10	≥1,120	≥6,280	≥860	≥600	≥8,880
Gender								
Female	<10	<10	<10	≥620	≥3,230	≥320	≥200	≥4,380
Male	<10	<10	<10	≥500	≥3,040	≥540	≥390	≥4,490
Ethnicity								
Hispanic/Latino	<10	<10	<10	≥160	≥610	≥170	≥150	≥1,100
American Indian or Alaska Native	<10	<10	<10	<10	≥30	<10	<10	≥40
Asian	<10	<10	<10	≥60	≥120	≥10	≥10	≥220
Black or African American	<10	<10	<10	≥290	≥2,820	≥420	≥330	≥3,880
Native Hawaiian or Other Pacific	<10	<10	<10	<10	<10	<10	<10	<10
White	<10	<10	<10	≥560	≥2,540	≥230	≥90	≥3,420
Two or More Races	<10	<10	<10	≥30	≥130	≥10	<10	≥190
Education Classification								
Regular	<10	<10	<10	≥980	≥5,860	≥750	≥500	≥8,100
Special	<10	<10	<10	≥10	≥210	≥100	≥90	≥420
Gifted	<10	<10	<10	≥120	≥200	<10	<10	≥340
Economic Status*								
Economically Disadvantaged	—	—	—	—	—	—	—	—
Not Economically Disadvantaged	—	—	—	—	—	—	—	—
English Learner Status								
Non-EL	<10	<10	<10	≥1,110	≥5,990	≥710	≥430	≥8,260
EL	<10	<10	<10	≥10	≥280	≥150	≥160	≥620
Migrant Status								
Non-migrant	<10	<10	<10	≥1,120	≥6,270	≥860	≥600	≥8,870
Migrant	<10	<10	<10	<10	<10	<10	<10	<10
Section 504 Status								
Non-section 504	<10	<10	<10	≥1,050	≥5,700	≥760	≥530	≥8,060
Section 504	<10	<10	<10	≥70	≥570	≥100	≥60	≥810
Homeless Status								
Not Homeless	<10	<10	<10	≥1,120	≥6,230	≥860	≥590	≥8,810
Homeless	<10	<10	<10	<10	≥40	<10	≥10	≥60
Military Affiliation								
Not Military Affiliated	<10	<10	<10	≥1,120	≥6,260	≥860	≥600	≥8,850
Military Affiliated	<10	<10	<10	<10	≥10	<10	<10	≥20
Foster Care Status								
Not in Foster Care	<10	<10	<10	≥1,120	≥6,280	≥860	≥600	≥8,870
Foster Care	<10	<10	<10	<10	<10	<10	<10	<10

* Economic status information is not available for the fall and summer administrations.

Table E.4 Percentage of Students taking the Fall 2020 Administration: English II

Group	Grade							
	6	7	8	9	10	11	12	Total
All Students	0.00	0.00	0.00	12.70	70.74	9.76	6.80	100
Gender								
Female	0.00	0.00	0.00	14.19	73.80	7.31	4.69	100
Male	0.00	0.00	0.00	11.24	67.74	12.15	8.86	100
Ethnicity								
Hispanic/Latino	0.00	0.00	0.00	14.92	55.15	15.91	14.01	100
American Indian or Alaska Native	0.00	0.00	0.00	10.64	78.72	6.38	4.26	100
Asian	0.00	0.00	0.00	30.04	57.85	6.73	5.38	100
Black or African American	0.00	0.00	0.00	7.70	72.69	10.93	8.68	100
Native Hawaiian or Other Pacific	0.00	0.00	0.00	16.67	66.67	0.00	16.67	100
White	0.00	0.00	0.00	16.37	74.18	6.83	2.63	100
Two or More Races	0.00	0.00	0.00	15.79	72.63	7.89	3.68	100
Education Classification								
Regular	0.00	0.00	0.00	12.13	72.32	9.31	6.24	100
Special	0.00	0.00	0.00	4.43	49.18	24.01	22.38	100
Gifted	0.00	0.00	0.00	36.42	60.40	2.60	0.58	100
Economic Status*								
Economically Disadvantaged	—	—	—	—	—	—	—	—
Not Economically Disadvantaged	—	—	—	—	—	—	—	—
English Learner Status								
Non-EL	0.00	0.00	0.00	14.30	75.37	8.21	2.11	100
EL	0.00	0.00	0.00	3.02	42.70	19.13	35.16	100
Migrant Status								
Non-migrant	0.00	0.00	0.00	13.45	72.59	8.68	5.28	100
Migrant	0.00	0.00	0.00	2.74	45.97	24.19	27.10	100
Section 504 Status								
Non-section 504	0.00	0.00	0.00	12.71	70.73	9.77	6.78	100
Section 504	0.00	0.00	0.00	0.00	71.43	0.00	28.57	100
Homeless Status								
Not Homeless	0.00	0.00	0.00	13.11	70.78	9.46	6.65	100
Homeless	0.00	0.00	0.00	8.68	70.29	12.71	8.31	100
Military Affiliation								
Not Military Affiliated	0.00	0.00	0.00	12.76	70.75	9.78	6.71	100
Military Affiliated	0.00	0.00	0.00	4.69	68.75	7.81	18.75	100
Foster Care Status								
Not in Foster Care	0.00	0.00	0.00	12.69	70.75	9.77	6.79	100
Foster Care	0.00	0.00	0.00	17.24	65.52	6.90	10.34	100

* Economic status information is not available for the fall and summer administrations.

Table E.5 Count of Students taking the Fall 2020 Administration: Algebra I

Group	Grade							
	6	7	8	9	10	11	12	Total
All Students	<10	<10	≥20	≥3,330	≥840	≥600	≥290	≥5,100
Gender								
Female	<10	<10	≥10	≥1,680	≥340	≥260	≥130	≥2,430
Male	<10	<10	≥10	≥1,650	≥500	≥330	≥160	≥2,670
Ethnicity								
Hispanic/Latino	<10	<10	<10	≥320	≥80	≥60	≥30	≥500
American Indian or Alaska Native	<10	<10	<10	≥10	<10	<10	<10	≥10
Asian	<10	<10	<10	≥80	<10	<10	<10	≥90
Black or African American	<10	<10	≥10	≥1,400	≥480	≥350	≥190	≥2,440
Native Hawaiian or Other Pacific	<10	<10	<10	<10	<10	<10	<10	<10
White	<10	<10	≥10	≥1,440	≥260	≥150	≥60	≥1,930
Two or More Races	<10	<10	<10	≥70	≥10	≥10	<10	≥90
Education Classification								
Regular	<10	<10	≥20	≥3,160	≥760	≥490	≥240	≥4,680
Special	<10	<10	<10	≥100	≥80	≥100	≥50	≥330
Gifted	<10	<10	<10	≥60	<10	<10	<10	≥80
Economic Status*								
Economically Disadvantaged	—	—	—	—	—	—	—	—
Not Economically Disadvantaged	—	—	—	—	—	—	—	—
English Learner Status								
Non-EL	<10	<10	≥20	≥3,220	≥760	≥530	≥260	≥4,810
EL	<10	<10	<10	≥110	≥80	≥60	≥30	≥280
Migrant Status								
Non-migrant	<10	<10	≥20	≥3,330	≥840	≥600	≥290	≥5,100
Migrant	<10	<10	<10	<10	<10	<10	<10	<10
Section 504 Status								
Non-section 504	<10	<10	≥10	≥3,040	≥720	≥520	≥250	≥4,560
Section 504	<10	<10	<10	≥290	≥120	≥80	≥40	≥540
Homeless Status								
Not Homeless	<10	<10	≥20	≥3,320	≥840	≥590	≥290	≥5,080
Homeless	<10	<10	<10	≥10	<10	<10	<10	≥20
Military Affiliation								
Not Military Affiliated	<10	<10	≥20	≥3,320	≥840	≥590	≥290	≥5,080
Military Affiliated	<10	<10	<10	≥10	<10	<10	<10	≥10
Foster Care Status								
Not in Foster Care	<10	<10	≥20	≥3,330	≥840	≥590	≥290	≥5,100
Foster Care	<10	<10	<10	<10	<10	<10	<10	<10

* Economic status information is not available for the fall and summer administrations.

Table E.6 Percentage of Students taking the Fall 2020 Administration: Algebra I

Group	Grade							
	6	7	8	9	10	11	12	Total
All Students	0.00	0.02	0.43	65.33	16.63	11.75	5.84	100
Gender								
Female	0.00	0.04	0.45	69.19	14.15	10.74	5.43	100
Male	0.00	0.00	0.41	61.83	18.88	12.67	6.21	100
Ethnicity								
Hispanic/Latino	0.00	0.00	0.20	63.24	16.40	13.44	6.72	100
American Indian or Alaska Native	0.00	0.00	0.00	55.56	22.22	11.11	11.11	100
Asian	0.00	0.00	0.00	85.26	6.32	4.21	4.21	100
Black or African American	0.00	0.04	0.41	57.29	19.78	14.55	7.93	100
Native Hawaiian or Other Pacific	0.00	0.00	0.00	66.67	33.33	0.00	0.00	100
White	0.00	0.00	0.57	74.66	13.47	8.05	3.25	100
Two or More Races	0.00	0.00	0.00	74.75	10.10	14.14	1.01	100
Education Classification								
Regular	0.00	0.00	0.43	67.64	16.24	10.52	5.17	100
Special	0.00	0.00	0.00	30.09	23.89	30.09	15.93	100
Gifted	0.00	1.22	2.44	79.27	8.54	6.10	2.44	100
Economic Status*								
Economically Disadvantaged	—	—	—	—	—	—	—	—
Not Economically Disadvantaged	—	—	—	—	—	—	—	—
English Learner Status								
Non-EL	0.00	0.02	0.46	66.97	15.94	11.15	5.46	100
EL	0.00	0.00	0.00	38.06	28.03	21.80	12.11	100
Migrant Status								
Non-migrant	0.00	0.02	0.43	65.35	16.62	11.76	5.82	100
Migrant	0.00	0.00	0.00	33.33	33.33	0.00	33.33	100
Section 504 Status								
Non-section 504	0.00	0.02	0.42	66.68	15.93	11.39	5.56	100
Section 504	0.00	0.00	0.55	53.97	22.55	14.79	8.13	100
Homeless Status								
Not Homeless	0.00	0.02	0.43	65.40	16.57	11.71	5.86	100
Homeless	0.00	0.00	0.00	52.00	28.00	20.00	0.00	100
Military Affiliation								
Not Military Affiliated	0.00	0.02	0.43	65.34	16.63	11.74	5.84	100
Military Affiliated	0.00	0.00	0.00	63.16	15.79	15.79	5.26	100
Foster Care Status								
Not in Foster Care	0.00	0.02	0.43	65.38	16.60	11.74	5.82	100
Foster Care	0.00	0.00	0.00	20.00	40.00	20.00	20.00	100

* Economic status information is not available for the fall and summer administrations.

Table E.7 Count of Students taking the Fall 2020 Administration: Geometry

Group	Grade							
	6	7	8	9	10	11	12	Total
All Students	<10	<10	<10	≥1,200	≥2,990	≥1,310	≥170	≥5,680
Gender								
Female	<10	<10	<10	≥670	≥1,570	≥680	≥70	≥3,000
Male	<10	<10	<10	≥530	≥1,420	≥620	≥100	≥2,680
Ethnicity								
Hispanic/Latino	<10	<10	<10	≥120	≥290	≥170	≥20	≥600
American Indian or Alaska Native	<10	<10	<10	<10	≥20	<10	<10	≥30
Asian	<10	<10	<10	≥50	≥50	≥20	<10	≥130
Black or African American	<10	<10	<10	≥340	≥1,340	≥740	≥120	≥2,550
Native Hawaiian or Other Pacific	<10	<10	<10	<10	<10	<10	<10	<10
White	<10	<10	<10	≥620	≥1,210	≥340	≥20	≥2,220
Two or More Races	<10	<10	<10	≥40	≥60	≥10	<10	≥120
Education Classification								
Regular	<10	<10	<10	≥1,040	≥2,820	≥1,250	≥150	≥5,280
Special	<10	<10	<10	<10	≥80	≥40	≥10	≥140
Gifted	<10	<10	<10	≥150	≥90	≥10	<10	≥260
Economic Status*								
Economically Disadvantaged	—	—	—	—	—	—	—	—
Not Economically Disadvantaged	—	—	—	—	—	—	—	—
English Learner Status								
Non-EL	<10	<10	<10	≥1,190	≥2,880	≥1,200	≥150	≥5,440
EL	<10	<10	<10	<10	≥110	≥100	≥20	≥240
Migrant Status								
Non-migrant	<10	<10	<10	≥1,200	≥2,990	≥1310	≥170	≥5,680
Migrant	<10	<10	<10	<10	<10	<10	<10	<10
Section 504 Status								
Non-section 504	<10	<10	<10	≥1,150	≥2,720	≥1,180	≥150	≥5,220
Section 504	<10	<10	<10	≥40	≥260	≥120	≥10	≥450
Homeless Status								
Not Homeless	<10	<10	<10	≥1,190	≥2,960	≥1,290	≥160	≥5,620
Homeless	<10	<10	<10	<10	≥20	≥10	<10	≥60
Military Affiliation								
Not Military Affiliated	<10	<10	<10	≥1,190	≥2,980	≥1,310	≥170	≥5,660
Military Affiliated	<10	<10	<10	<10	<10	<10	<10	≥10
Foster Care Status								
Not in Foster Care	<10	<10	<10	≥1,200	≥2,990	≥1,310	≥170	≥5,680
Foster Care	<10	<10	<10	<10	<10	<10	<10	<10

* Economic status information is not available for the fall and summer administrations.

Table E.8 Percentage of Students taking the Fall 2020 Administration: Geometry

Group	Grade							
	6	7	8	9	10	11	12	Total
All Students	0.00	0.00	0.02	21.17	52.72	23.07	3.02	100
Gender								
Female	0.00	0.00	0.00	22.37	52.50	22.80	2.33	100
Male	0.00	0.00	0.04	19.83	52.96	23.37	3.80	100
Ethnicity								
Hispanic/Latino	0.00	0.00	0.00	20.53	48.11	27.91	3.45	100
American Indian or Alaska Native	0.00	0.00	0.00	12.90	67.74	19.35	0.00	100
Asian	0.00	0.00	0.00	41.30	39.13	18.12	1.45	100
Black or African American	0.00	0.00	0.00	13.51	52.55	29.25	4.70	100
Native Hawaiian or Other Pacific	0.00	0.00	0.00	16.67	83.33	0.00	0.00	100
White	0.00	0.00	0.05	28.32	54.71	15.62	1.31	100
Two or More Races	0.00	0.00	0.00	33.59	53.13	13.28	0.00	100
Education Classification								
Regular	0.00	0.00	0.02	19.78	53.48	23.75	2.97	100
Special	0.00	0.00	0.00	4.93	57.04	28.17	9.86	100
Gifted	0.00	0.00	0.00	58.24	34.87	6.51	0.38	100
Economic Status*								
Economically Disadvantaged	—	—	—	—	—	—	—	—
Not Economically Disadvantaged	—	—	—	—	—	—	—	—
English Learner Status								
Non-EL	0.00	0.00	0.02	21.99	53.01	22.19	2.79	100
EL	0.00	0.00	0.00	2.88	46.09	42.80	8.23	100
Migrant Status								
Non-migrant	0.00	0.00	0.02	21.19	52.73	23.07	2.99	100
Migrant	0.00	0.00	0.00	0.00	40.00	20.00	40.00	100
Section 504 Status								
Non-section 504	0.00	0.00	0.02	22.09	52.19	22.66	3.04	100
Section 504	0.00	0.00	0.00	10.70	58.73	27.73	2.84	100
Homeless Status								
Not Homeless	0.00	0.00	0.02	21.24	52.76	22.98	3.00	100
Homeless	0.00	0.00	0.00	15.00	48.33	31.67	5.00	100
Military Affiliation								
Not Military Affiliated	0.00	0.00	0.02	21.10	52.73	23.13	3.03	100
Military Affiliated	0.00	0.00	0.00	44.44	50.00	5.56	0.00	100
Foster Care Status								
Not in Foster Care	0.00	0.00	0.02	21.18	52.71	23.08	3.01	100
Foster Care	0.00	0.00	0.00	0.00	66.67	0.00	33.33	100

* Economic status information is not available for the fall and summer administrations.

Table E.9 Count of Students taking the Summer 2021 Administration: English I

Group	Grade							
	6	7	8	9	10	11	12	Total
All Students	<10	<10	<10	≥1,930	≥420	≥200	≥10	≥2,580
Gender								
Female	<10	<10	<10	≥640	≥130	≥70	<10	≥860
Male	<10	<10	<10	≥1,290	≥280	≥120	<10	≥1,710
Ethnicity								
Hispanic/Latino	<10	<10	<10	≥180	≥70	≥30	<10	≥290
American Indian or Alaska Native	<10	<10	<10	<10	<10	<10	<10	≥10
Asian	<10	<10	<10	≥10	<10	<10	<10	≥20
Black or African American	<10	<10	<10	≥1,300	≥260	≥140	≥10	≥1,720
Native Hawaiian or Other Pacific	<10	<10	<10	<10	<10	<10	<10	<10
White	<10	<10	<10	≥400	≥60	≥30	<10	≥500
Two or More Races	<10	<10	<10	≥30	<10	<10	<10	≥30
Education Classification								
Regular	<10	<10	<10	≥1,430	≥330	≥140	<10	≥1,920
Special	<10	<10	<10	≥480	≥80	≥60	<10	≥630
Gifted	<10	<10	<10	≥10	<10	<10	<10	≥10
Economic Status*								
Economically Disadvantaged	—	—	—	—	—	—	—	—
Not Economically Disadvantaged	—	—	—	—	—	—	—	—
English Learner Status								
Non-EL	<10	<10	<10	≥1,790	≥350	≥170	≥10	≥2,330
EL	<10	<10	<10	≥140	≥70	≥30	<10	≥250
Migrant Status								
Non-migrant	<10	<10	<10	≥1,930	≥420	≥200	≥10	≥2,570
Migrant	<10	<10	<10	<10	<10	<10	<10	<10
Section 504 Status								
Non-section 504	<10	<10	<10	≥1,620	≥340	≥160	≥10	≥2,150
Section 504	<10	<10	<10	≥310	≥70	≥30	<10	≥420
Homeless Status								
Not Homeless	<10	<10	<10	≥1,900	≥410	≥190	≥10	≥2,520
Homeless	<10	<10	<10	≥30	<10	<10	<10	≥50
Military Affiliation								
Not Military Affiliated	<10	<10	<10	≥1,930	≥410	≥200	≥10	≥2,560
Military Affiliated	<10	<10	<10	<10	<10	<10	<10	≥10
Foster Care Status								
Not in Foster Care	<10	<10	<10	≥1,930	≥420	≥200	≥10	≥2,570
Foster Care	<10	<10	<10	<10	<10	<10	<10	≥10

* Economic status information is not available for the fall and summer administrations.

Table E.10 Percentage of Students taking the Summer 2021 Administration: English I

Group	Grade							
	6	7	8	9	10	11	12	Total
All Students	0.00	0.00	0.04	75.13	16.35	7.98	0.50	100
Gender								
Female	0.00	0.00	0.12	74.83	15.59	9.01	0.46	100
Male	0.00	0.00	0.00	75.28	16.73	7.46	0.52	100
Ethnicity								
Hispanic/Latino	0.00	0.00	0.00	62.07	26.55	10.34	1.03	100
American Indian or Alaska Native	0.00	0.00	0.00	60.00	40.00	0.00	0.00	100
Asian	0.00	0.00	0.00	76.19	14.29	9.52	0.00	100
Black or African American	0.00	0.00	0.00	75.78	15.45	8.19	0.58	100
Native Hawaiian or Other Pacific								
White	0.00	0.00	0.20	80.08	13.75	5.98	0.00	100
Two or More Races	0.00	0.00	0.00	83.33	8.33	8.33	0.00	100
Education Classification								
Regular	0.00	0.00	0.05	74.74	17.31	7.54	0.36	100
Special	0.00	0.00	0.00	76.06	13.62	9.39	0.94	100
Gifted	0.00	0.00	0.00	83.33	11.11	5.56	0.00	100
Economic Status*								
Economically Disadvantaged	—	—	—	—	—	—	—	—
Not Economically Disadvantaged	—	—	—	—	—	—	—	—
English Learner Status								
Non-EL	0.00	0.00	0.04	77.05	15.06	7.42	0.43	100
EL	0.00	0.00	0.00	57.20	28.40	13.20	1.20	100
Migrant Status								
Non-migrant	0.00	0.00	0.04	75.14	16.32	8.00	0.51	100
Migrant	0.00	0.00	0.00	71.43	28.57	0.00	0.00	100
Section 504 Status								
Non-section 504	0.00	0.00	0.00	75.43	16.21	7.80	0.56	100
Section 504	0.00	0.00	0.23	73.60	17.06	8.88	0.23	100
Homeless Status								
Not Homeless	0.00	0.00	0.04	75.25	16.44	7.80	0.48	100
Homeless	0.00	0.00	0.00	69.64	12.50	16.07	1.79	100
Military Affiliation								
Not Military Affiliated	0.00	0.00	0.04	75.17	16.27	8.02	0.51	100
Military Affiliated	0.00	0.00	0.00	66.67	33.33	0.00	0.00	100
Foster Care Status								
Not in Foster Care	0.00	0.00	0.04	75.15	16.34	7.97	0.51	100
Foster Care	0.00	0.00	0.00	70.00	20.00	10.00	0.00	100

* Economic status information is not available for the fall and summer administrations.

Table E.11 Count of Students taking the Summer 2021 Administration: English II

Group	Grade							
	6	7	8	9	10	11	12	Total
All Students	<10	<10	<10	≥120	≥2,140	≥360	≥10	≥2,640
Gender								
Female	<10	<10	<10	≥30	≥780	≥110	<10	≥930
Male	<10	<10	<10	≥80	≥1,360	≥240	≥10	≥1,700
Ethnicity								
Hispanic/Latino	<10	<10	<10	≥10	≥120	≥40	<10	≥180
American Indian or Alaska Native	<10	<10	<10	<10	≥10	<10	<10	≥10
Asian	<10	<10	<10	<10	≥10	<10	<10	≥10
Black or African American	<10	<10	<10	≥80	≥1,490	≥220	≥10	≥1,810
Native Hawaiian or Other Pacific	<10	<10	<10	<10	<10	<10	<10	<10
White	<10	<10	<10	≥20	≥460	≥80	<10	≥570
Two or More Races	<10	<10	<10	<10	<10	<10	<10	≥30
Education Classification								
Regular	<10	<10	<10	≥90	≥1,670	≥250	≥10	≥2,040
Special	<10	<10	<10	≥20	≥440	≥100	<10	≥570
Gifted	<10	<10	<10	<10	≥20	<10	<10	≥20
Economic Status*								
Economically Disadvantaged	—	—	—	—	—	—	—	—
Not Economically Disadvantaged	—	—	—	—	—	—	—	—
English Learner Status								
Non-EL	<10	<10	<10	≥110	≥2,030	≥320	≥10	≥2,470
EL	<10	<10	<10	≥10	≥110	≥40	<10	≥160
Migrant Status								
Non-migrant	<10	<10	<10	≥120	≥2,140	≥360	≥10	≥2,640
Migrant	<10	<10	<10	<10	<10	<10	<10	<10
Section 504 Status								
Non-section 504	<10	<10	<10	≥100	≥1,790	≥290	≥10	≥2,200
Section 504	<10	<10	<10	≥20	≥340	≥60	<10	≥430
Homeless Status								
Not Homeless	<10	<10	<10	≥110	≥2,100	≥350	≥10	≥2,580
Homeless	<10	<10	<10	<10	≥40	≥10	<10	≥50
Military Affiliation								
Not Military Affiliated	<10	<10	<10	≥120	≥2,120	≥360	≥10	≥2,610
Military Affiliated	<10	<10	<10	<10	≥20	<10	<10	≥20
Foster Care Status								
Not in Foster Care	<10	<10	<10	≥120	≥2,130	≥360	≥10	≥2,620
Foster Care	<10	<10	<10	<10	≥10	<10	<10	≥10

* Economic status information is not available for the fall and summer administrations.

Table E.12 Percentage of Students taking the Summer 2021 Administration: English II

Group	Grade							
	6	7	8	9	10	11	12	Total
All Students	0.00	0.00	0.00	4.62	81.15	13.70	0.53	100
Gender								
Female	0.00	0.00	0.00	4.06	83.24	12.27	0.43	100
Male	0.00	0.00	0.00	4.93	80.00	14.49	0.59	100
Ethnicity								
Hispanic/Latino	0.00	0.00	0.00	7.03	69.19	22.16	1.62	100
American Indian or Alaska Native	0.00	0.00	0.00	0.00	94.12	5.88	0.00	100
Asian	0.00	0.00	0.00	11.76	76.47	11.76	0.00	100
Black or African American	0.00	0.00	0.00	4.40	82.41	12.59	0.60	100
Native Hawaiian or Other Pacific	0.00	0.00	0.00	0.00	100.00	0.00	0.00	100
White	0.00	0.00	0.00	4.21	81.05	14.74	0.00	100
Two or More Races	0.00	0.00	0.00	9.09	75.76	15.15	0.00	100
Education Classification								
Regular	0.00	0.00	0.00	4.75	82.21	12.45	0.59	100
Special	0.00	0.00	0.00	4.17	76.87	18.61	0.35	100
Gifted	0.00	0.00	0.00	3.70	92.59	3.70	0.00	100
Economic Status*								
Economically Disadvantaged	—	—	—	—	—	—	—	—
Not Economically Disadvantaged	—	—	—	—	—	—	—	—
English Learner Status								
Non-EL	0.00	0.00	0.00	4.52	82.04	12.99	0.44	100
EL	0.00	0.00	0.00	6.10	67.68	24.39	1.83	100
Migrant Status								
Non-migrant	0.00	0.00	0.00	4.62	81.15	13.70	0.53	100
Migrant								
Section 504 Status								
Non-section 504	0.00	0.00	0.00	4.54	81.45	13.47	0.54	100
Section 504	0.00	0.00	0.00	5.03	79.63	14.87	0.46	100
Homeless Status								
Not Homeless	0.00	0.00	0.00	4.57	81.31	13.62	0.50	100
Homeless	0.00	0.00	0.00	6.90	74.14	17.24	1.72	100
Military Affiliation								
Not Military Affiliated	0.00	0.00	0.00	4.62	81.06	13.78	0.53	100
Military Affiliated	0.00	0.00	0.00	4.35	91.30	4.35	0.00	100
Foster Care Status								
Not in Foster Care	0.00	0.00	0.00	4.60	81.17	13.73	0.49	100
Foster Care	0.00	0.00	0.00	7.69	76.92	7.69	7.69	100

* Economic status information is not available for the fall and summer administrations.

Table E.13 Count of Students taking the Summer 2021 Administration: Algebra I

Group	Grade							
	6	7	8	9	10	11	12	Total
All Students	<10	<10	≥90	≥2,590	≥670	≥150	≥20	≥3,540
Gender								
Female	<10	<10	≥40	≥1,150	≥270	≥70	<10	≥1,550
Male	<10	<10	≥50	≥1,440	≥390	≥80	≥10	≥1,980
Ethnicity								
Hispanic/Latino	<10	<10	<10	≥190	≥50	≥10	<10	≥270
American Indian or Alaska Native	<10	<10	<10	≥10	<10	<10	<10	≥10
Asian	<10	<10	<10	<10	<10	<10	<10	≥10
Black or African American	<10	<10	≥10	≥1,760	≥450	≥110	≥10	≥2,360
Native Hawaiian or Other Pacific	<10	<10	<10	<10	<10	<10	<10	<10
White	<10	<10	<10	≥550	≥140	≥30	<10	≥790
Two or More Races	<10	<10	<10	≥50	<10	<10	<10	≥70
Education Classification								
Regular	<10	<10	≥70	≥2,080	≥520	≥120	≥10	≥2,820
Special	<10	<10	<10	≥480	≥140	≥30	<10	≥660
Gifted	<10	<10	≥10	≥30	<10	<10	<10	≥50
Economic Status*								
Economically Disadvantaged	—	—	—	—	—	—	—	—
Not Economically Disadvantaged	—	—	—	—	—	—	—	—
English Learner Status								
Non-EL	<10	<10	≥90	≥2,450	≥630	≥140	≥20	≥3,350
EL	<10	<10	<10	≥130	≥40	≥10	<10	≥180
Migrant Status								
Non-migrant	<10	<10	≥90	≥2,580	≥670	≥150	≥20	≥3,530
Migrant	<10	<10	<10	<10	<10	<10	<10	<10
Section 504 Status								
Non-section 504	<10	<10	≥80	≥2,180	≥560	≥130	≥20	≥2,990
Section 504	<10	<10	<10	≥400	≥100	≥20	<10	≥540
Homeless Status								
Not Homeless	<10	<10	≥90	≥2,540	≥660	≥150	≥20	≥3,470
Homeless	<10	<10	<10	≥40	≥10	<10	<10	≥60
Military Affiliation								
Not Military Affiliated	<10	<10	≥90	≥2,570	≥670	≥150	≥20	≥3,520
Military Affiliated	<10	<10	<10	≥10	<10	<10	<10	≥10
Foster Care Status								
Not in Foster Care	<10	<10	≥90	≥2,580	≥660	≥150	≥20	≥3,520
Foster Care	<10	<10	<10	<10	<10	<10	<10	≥10

* Economic status information is not available for the fall and summer administrations.

Table E.14 Percentage of Students taking the Summer 2021 Administration: Algebra I

Group	Grade							
	6	7	8	9	10	11	12	Total
All Students	0.00	0.00	2.65	73.14	19.06	4.49	0.65	100
Gender								
Female	0.00	0.00	2.70	73.95	17.94	4.95	0.45	100
Male	0.00	0.00	2.62	72.51	19.94	4.13	0.81	100
Ethnicity								
Hispanic/Latino	0.00	0.00	1.48	73.33	20.00	4.07	1.11	100
American Indian or Alaska Native	0.00	0.00	0.00	68.75	31.25	0.00	0.00	100
Asian	0.00	0.00	47.37	47.37	5.26	0.00	0.00	100
Black or African American	0.00	0.00	0.72	74.39	19.40	4.73	0.76	100
Native Hawaiian or Other Pacific	0.00	0.00	0.00	100.00	0.00	0.00	0.00	100
White	0.00	0.00	7.76	69.34	18.65	4.13	0.13	100
Two or More Races	0.00	0.00	2.86	81.43	10.00	4.29	1.43	100
Education Classification								
Regular	0.00	0.00	2.58	73.60	18.58	4.56	0.67	100
Special	0.00	0.00	0.45	72.40	22.32	4.52	0.30	100
Gifted	0.00	0.00	34.62	57.69	3.85	0.00	3.85	100
Economic Status*								
Economically Disadvantaged	—	—	—	—	—	—	—	—
Not Economically Disadvantaged	—	—	—	—	—	—	—	—
English Learner Status								
Non-EL	0.00	0.00	2.80	73.29	18.87	4.41	0.63	100
EL	0.00	0.00	0.00	70.59	22.46	5.88	1.07	100
Migrant Status								
Non-migrant	0.00	0.00	2.63	73.18	19.04	4.50	0.65	100
Migrant	0.00	0.00	16.67	50.00	33.33	0.00	0.00	100
Section 504 Status								
Non-section 504	0.00	0.00	2.97	73.04	18.94	4.34	0.70	100
Section 504	0.00	0.00	0.91	73.72	19.71	5.29	0.36	100
Homeless Status								
Not Homeless	0.00	0.00	2.70	73.15	18.98	4.51	0.66	100
Homeless	0.00	0.00	0.00	73.02	23.81	3.17	0.00	100
Military Affiliation								
Not Military Affiliated	0.00	0.00	2.67	73.09	19.09	4.51	0.65	100
Military Affiliated	0.00	0.00	0.00	86.67	13.33	0.00	0.00	100
Foster Care Status								
Not in Foster Care	0.00	0.00	2.66	73.25	18.93	4.51	0.65	100
Foster Care	0.00	0.00	0.00	41.67	58.33	0.00	0.00	100

* Economic status information is not available for the fall and summer administrations.

Table E.15 Count of Students taking the Summer 2021 Administration: Geometry

Group	Grade							
	6	7	8	9	10	11	12	Total
All Students	<10	<10	<10	≥50	≥760	≥230	<10	≥1,050
Gender								
Female	<10	<10	<10	≥30	≥400	≥140	<10	≥570
Male	<10	<10	<10	≥20	≥360	≥80	<10	≥470
Ethnicity								
Hispanic/Latino	<10	<10	<10	<10	≥40	≥10	<10	≥60
American Indian or Alaska Native	<10	<10	<10	<10	<10	<10	<10	<10
Asian	<10	<10	<10	<10	<10	<10	<10	<10
Black or African American	<10	<10	<10	≥40	≥540	≥170	<10	≥760
Native Hawaiian or Other Pacific	<10	<10	<10	<10	<10	<10	<10	<10
White	<10	<10	<10	<10	≥150	≥40	<10	≥200
Two or More Races	<10	<10	<10	<10	≥10	<10	<10	≥10
Education Classification								
Regular	<10	<10	<10	≥40	≥650	≥190	<10	≥900
Special	<10	<10	<10	<10	≥90	≥30	<10	≥130
Gifted	<10	<10	<10	<10	<10	<10	<10	≥10
Economic Status*								
Economically Disadvantaged	—	—	—	—	—	—	—	—
Not Economically Disadvantaged	—	—	—	—	—	—	—	—
English Learner Status								
Non-EL	<10	<10	<10	≥50	≥740	≥210	<10	≥1,020
EL	<10	<10	<10	<10	≥20	≥10	<10	≥30
Migrant Status								
Non-migrant	<10	<10	<10	≥50	≥760	≥220	<10	≥1,050
Migrant	<10	<10	<10	<10	<10	<10	<10	<10
Section 504 Status								
Non-section 504	<10	<10	<10	≥50	≥670	≥200	<10	≥930
Section 504	<10	<10	<10	<10	≥90	≥20	<10	≥110
Homeless Status								
Not Homeless	<10	<10	<10	≥50	≥750	≥220	<10	≥1,030
Homeless	<10	<10	<10	<10	≥10	<10	<10	≥10
Military Affiliation								
Not Military Affiliated	<10	<10	<10	≥50	≥750	≥220	<10	≥1,050
Military Affiliated	<10	<10	<10	<10	<10	<10	<10	<10
Foster Care Status								
Not in Foster Care	<10	<10	<10	≥50	≥760	≥230	<10	≥1,050
Foster Care	<10	<10	<10	<10	<10	<10	<10	<10

* Economic status information is not available for the fall and summer administrations.

Table E.16 Percentage of Students taking the Summer 2021 Administration: Geometry

Group	Grade							
	6	7	8	9	10	11	12	Total
All Students	0.00	0.00	0.00	5.39	72.21	21.74	0.66	100
Gender								
Female	0.00	0.00	0.00	5.70	69.26	24.53	0.52	100
Male	0.00	0.00	0.00	5.01	75.78	18.37	0.84	100
Ethnicity								
Hispanic/Latino	0.00	0.00	0.00	6.06	69.70	21.21	3.03	100
American Indian or Alaska Native	0.00	0.00	0.00	0.00	100.00	0.00	0.00	100
Asian	0.00	0.00	0.00	12.50	75.00	12.50	0.00	100
Black or African American	0.00	0.00	0.00	5.37	71.60	22.38	0.65	100
Native Hawaiian or Other Pacific								
White	0.00	0.00	0.00	4.48	75.62	19.90	0.00	100
Two or More Races	0.00	0.00	0.00	11.11	66.67	22.22	0.00	100
Education Classification								
Regular	0.00	0.00	0.00	5.09	72.79	21.35	0.77	100
Special	0.00	0.00	0.00	4.38	72.26	23.36	0.00	100
Gifted	0.00	0.00	0.00	29.41	41.18	29.41	0.00	100
Economic Status*								
Economically Disadvantaged	—	—	—	—	—	—	—	—
Not Economically Disadvantaged	—	—	—	—	—	—	—	—
English Learner Status								
Non-EL	0.00	0.00	0.00	5.39	72.67	21.45	0.49	100
EL	0.00	0.00	0.00	5.41	59.46	29.73	5.41	100
Migrant Status								
Non-migrant	0.00	0.00	0.00	5.40	72.25	21.69	0.66	100
Migrant	0.00	0.00	0.00	0.00	50.00	50.00	0.00	100
Section 504 Status								
Non-section 504	0.00	0.00	0.00	5.43	71.67	22.26	0.64	100
Section 504	0.00	0.00	0.00	5.04	76.47	17.65	0.84	100
Homeless Status								
Not Homeless	0.00	0.00	0.00	5.39	72.38	21.66	0.58	100
Homeless	0.00	0.00	0.00	5.26	63.16	26.32	5.26	100
Military Affiliation								
Not Military Affiliated	0.00	0.00	0.00	5.43	72.10	21.81	0.67	100
Military Affiliated	0.00	0.00	0.00	0.00	87.50	12.50	0.00	100
Foster Care Status								
Not in Foster Care	0.00	0.00	0.00	5.40	72.16	21.78	0.66	100
Foster Care	0.00	0.00	0.00	0.00	100.00	0.00	0.00	100

* Economic status information is not available for the fall and summer administrations.

References

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

Beimers, J. N., Way, W. D., McClarty, K. L., & Miles, J. A. (2012, January). Evidence based standard setting: Establishing cut scores by integrating research evidence with expert content judgments. Austin, TX: Pearson. Retrieved from http://researchnetwork.pearson.com/wpcontent/uploads/Bulletin21_Evidence_Based_Standard_Setting.pdf

Cai, L., Thissen, D., & du Toit, S. H. C. (2011). IRTPRO for Windows [Computer software]. Lincolnwood, IL: Scientific Software International.

Camilli, G., & Shepard, A. L. (1994). *Methods for identifying biased test items*. Thousand Oaks, CA: Sage Publication.

Center for Assessment. (2017, June). *LEAP 2017: English language arts -grade 6 summary – comparability with PARCC performance standards* (Memorandum). Dove, NH.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Earlbaum Associates.

Crocker, L., & Algina, J. (1986). *Introduction to classical and modern test theory*. Belmont, CA: Wadsworth.

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

Data Recognition Corporation. (2016). *Interpretive guide: Grades 3–8 ELA and math.* Maple Grove, MN: Author.

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

Educational Testing Service, Pearson, & Measured Progress. (2016). *Final technical report for 2015 administration*. PARCC. Retrieved from https://eric.ed.gov/?q=source%3a%22Partnership+for+Assessment+of+Readiness+for+College+and+Careers%22&id=ED599097

Green, D. R. (1975). *Procedures for assessing bias in achievement tests*. Paper presented at the National Institute of Education Conference on Test Bias, Annapolis, MD.

Hambleton, R. K., & Swaminathan, H. (1985). *Item response theory: Principles and applications*. Hingham, MA: Kluwer-Nijhoff Publishing.

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

Kim, S., & Kolen, M. (2004). STUIRT: A computer program for scale transformation under unidimensional item repose theory models (Version 1.0) [Computer software]. Iowa City, IA: University of Iowa.

Kolen, M. J., & Brennan, R. L. (2014). *Test equating, scaling, and linking*. New York, NY: Springer-Verlag.

Livingston, S. A., & Lewis, C. (1995). Estimating the consistency and accuracy of classifications based on test scores. *Journal of Educational Measurement*, *32(2)*, 179–197.

Mantel, N., & Haenszel, W. (1959). Statistical aspects of the analysis of data from retrospective studies of disease. *Journal of the National Cancer Institute*, *22(4)*, 719–748.

Muraki, E. (1992). A generalized partial credit model: Application of an EM algorithm. *Applied Psychological Measurement*, *16(2)*, 159–176.

Pearson. (2015). *Performance level setting technical report*. PARCC. Retrieved from https://eric.ed.gov/?q=source%3a%22Partnership+for+Assessment+of+Readiness+for+College+and+Careers%22&pg=2&id=ED599257

Pearson. (2017). *PARCC: Final technical report for 2016 administration.* PARCC. Retrieved from https://eric.ed.gov/?q=PARCC+Final+Technical+Report&id=ED599197

Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, *70*(1), 41–55.

Schumacker, R. E. (1996). Disattenuating correlation coefficients. *Rasch Measurement Transactions*, *10(1)*, 479.

Stocking, M. L., & Lord, F. M. (1983). Developing a common metric in item response theory. *Applied Psychological Measurement*, *7(2)*, 201–210.

Thompson, S., & Thurlow, M. (2002). *Universally designed assessments: Better tests for everyone!* (Policy Directions No. 14). Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes. Retrieved from http://www.cehd.umn.edu/NCEO/OnlinePUbs/Policy14.htm

Zwick, R., Donoghue, J. R., & Grima, A. (1993). Assessment of differential item functioning for performance tasks. *Journal of Educational Measurement*, *30(3)*, 233–251.