PAUSD Validation Test Research Brief

WestEd

May 2025

WestEd is a nonpartisan, nonprofit organization that aims to improve the lives of children and adults at all ages of learning and development. We do this by addressing challenges in education and human development, increasing opportunity, and helping build communities where all can thrive. WestEd staff conduct and apply research, provide technical assistance, and support professional learning. We work with early learning educators, classroom teachers, local and state leaders, and policymakers at all levels of government.

For more information, visit WestEd.org. For regular updates on research, free resources, solutions, and job postings from WestEd, subscribe to the E-Bulletin, our semimonthly e-newsletter, at WestEd.org/subscribe.

Table of Contents

Executive Summary

In the 2023/24 academic year, Palo Alto Unified School District (PAUSD) aimed to update three existing assessments that have been used to identify students who are eligible to participate in the Middle School Math Acceleration Process. PAUSD requested that WestEd assessment experts and psychometricians support the development of these three assessments, commonly referred to as validation tests. Specifically, these three tests are aligned with the core courses in grade 6, grade 7, and grade 8, respectively, and are used to determine whether a student has demonstrated sufficient proficiency in the course subject matter to bypass the course and accelerate to the next level. The three courses—Foundations for grade 6, Concepts for grade 7, and Algebra for grade 8—are a compacted middle school math program (four years of math in three years). Students who are eligible to accelerate beyond this already compacted program will, in effect, cover five years of math content in three academic years.

The process of updating the validation tests included blueprint and item development, item tryouts, a cut-score study, and operational data analyses. These procedures resulted in final cut scores that were applied to the spring 2024 operational administrations.

- In Foundations, the qualifying score was 70 percent, and 36 percent of students who took the test qualified for acceleration.
- In Concepts, the qualifying score was 60 percent, and 23 percent of students who took the test qualified for acceleration.
- In Algebra, the qualifying score was 70 percent, and 20 percent of students who took the test qualified for acceleration.

This research brief outlines the research and procedures used to develop and administer the assessments and to identify the appropriate cut scores to use as passing thresholds. The work was done in collaboration among PAUSD instructional leads, PAUSD district personnel, WestEd assessment and content experts, and WestEd psychometricians.

Introduction

Background

Palo Alto Unified School District (PAUSD) has been administering validation tests as part of its Math Validation Process to determine student eligibility for acceleration through its middle school math program. In the 2023/24 academic year, PAUSD sought to update these tests, with technical assistance provided by WestEd. Specifically, PAUSD sought to update the Foundations, Concepts, and Algebra assessments that are used to determine student eligibility to accelerate through the grade 6, grade 7, and grade 8 courses, respectively.

For this process, WestEd has provided technical assistance in the form of assessment development services such as item writing, blueprints, and test maps, as well as psychometric services in the form of item and test analysis and cut-score establishment. The work was done in collaboration with PAUSD middle school math instructional leads, PAUSD district personnel, WestEd assessment and content experts, and WestEd psychometricians.

Validation Test Goals

The goal of the validation tests is to determine whether a student is demonstrating enough proficiency with the content of a course to bypass the course and accelerate to the next level in the middle school math program. Three validation tests have been developed: the Foundations test, which is used to determine whether a student qualifies to bypass the grade 6 Foundations math course; the Concepts test, which is used to determine whether a student qualifies to bypass the grade 7 Concepts math course; and the Algebra test, which is used to determine whether a student qualifies to bypass the grade 8 Algebra math course.

Validation Test Research

The research outlined in this brief aims to support the development of the validation test forms and to set appropriate cut scores that reflect readiness for the next level in the math program. The research involved numerous studies, including preliminary item tryouts, operational data analyses, and a cut-score setting study. Each of these studies is discussed in the relevant section of the document.

Test Development

Test Blueprints

Previous versions of the validation tests included two parts that were administered at separate times. Part 1 included only multiple-choice (MC) items, and students had to pass Part 1 to qualify for Part 2. Part 2 was composed solely of free-response questions (FRQs). Only students who received a qualifying score on Part 2 were eligible for acceleration.

The validation tests that were developed for spring 2024 are administered in a single testing session. The tests are composed of a combination of a subset of the previously used FRQs (which were reviewed and validated by WestEd staff) and newly developed MC items. The blueprints include a percentage of coverage for each domain in each course, with a possible point range for the domain. The total point ranges include the numbers of points from the FRQs and the MC items. Not all standards are covered within the blueprints. This ensures that the assessment can be completed within a reasonable amount of time, while providing detailed information on the assessed content.

Table 1, Table 2, and Table 3 present the percentages of points and the ranges of points per domain in the Foundations, Concepts, and Algebra validation tests, respectively.

Table 1. Percentage of Points in Each Domain in the Foundations Validation Test

Domain	Total Percentage of Points	Range of Points
Number Sense (NS)	5	3–5
Statistics and Probability (SP)	10	6–9
Equations and Expressions (EE)	35	25–30
Ratios and Proportions (RP)	35	25–30
Geometry (G)	15	12–14

Table 2. Percentage of Points in Each Domain in the Concepts Validation Test

Domain	Total Percentage of Points	Range of Points
Number Sense (NS)	5	3–5
Statistics and Probability (SP)	10	5–10
Equations and Expressions (EE)	35	25–30
Ratios and Proportions (RP)	15	12–15
Geometry (G)	35	25–30

Table 3. Percentage of Points in Each Domain in the Algebra Validation Test

Domain	Total Percentage of Points	Range of Points
Equations and Expressions (EE)	30	23–27
Functions (F) INCLUDES • Grade 8 Functions • Functions—Building Functions (F.BF) • Functions—Interpreting Functions (F.IF) • Functions—Linear, Quadratic, and Exponential Models (F.LE)	30	23–27
ALGEBRA (A) INCLUDES • Algebra—Seeing Structure in Expressions (A.SSE) • Algebra—Arithmetic with Polynomials and Rational Expressions (A.APR) • Algebra—Creating Equations (A.CED)	30	23–27
NUMBERS (N) INCLUDES • Numbers—The Real Number System (N.RN) • Numbers—Quantities (N.Q)	10	8–10

Item Development

All MC items were developed by WestEd. Item writers were trained in general item-writing techniques and in writing guidelines that were specific to the needs of the validation tests. WestEd reviewed submissions from item writers, determining whether each item should move forward to the item review and revision process.

All MC items underwent a series of internal reviews within WestEd to ensure that they met assessment industry standards and that they adhered to the Standards for Educational and Psychological Testing (AERA et al., 2014). WestEd implemented several processes and quality-control procedures to support the development of MC items that:

- measure the subject-specific standards with a consistent interpretation of the standards across item writers, content specialists, and reviewers;
- include only one correct answer;
- have four answer choices that are balanced grammatically and structurally;
- have plausible but incorrect options;
- do not clue other items;
- are developmentally appropriate for the given grade, through engagement of content specialists and reviewers familiar with the knowledge, skills, and abilities of students at the tested grades;
- maximize accessibility for students through the application of universal design principles;
- are clear, concise, and free of idiomatic expressions and grammatical errors;
- exclude bias, stereotyping information, and sensitive content; and
- exhibit strong alignment to the intended construct.

Items were submitted to PAUSD staff for review in two batches: one for calibration and one that included the remainder of the items, along with any revisions necessary to the calibration items. PAUSD staff reviewed the items and provided feedback to WestEd content specialists within Smartsheet, a secure project management–tracking platform. Feedback and suggestions by PAUSD were incorporated into the items, as necessary.

Guidelines for Fairness and Bias

Items were developed using universal design principles to ensure that the items are accessible to the widest range of students, spanning a range of complexity. These principles include:

- maximum legibility—capability of being deciphered at ease;
- readability and comprehensibility—being written at grade level in terms of content or subject matter, vocabulary, and readability; and
- prior knowledge—expecting appropriate grade-level knowledge is acceptable.

Additionally, bias and sensitivity considerations ensured that items:

- do not advantage or disadvantage any group of students;
- are free of stereotypes (e.g., cultural, racial, gender); and
- avoid sensitive topics or contexts (e.g., politics, religion, death, illness, pandemics).

Free-Response Questions and Rubrics

Previously used FRQs were reviewed and validated, and a subset of questions were selected for use on the spring 2024 test. The complexity of each FRQ and previous scoring information were used by WestEd to make recommendations regarding the maximum point value of each FRQ, which are referred to as "Max Points" in tables that appear later in this research brief. A rubric to guide the scoring of each FRQ was developed to describe the possible score points ranging from zero points to the "Max Points." Rubrics were reviewed and accepted by the middle school math instructional leads.

Item Tryouts and Psychometric Review

Item tryouts were conducted for each test with up to 45 MC items. The forms were given to on-grade-level students in the course meant to be accelerated; for example, the Concepts form was given to students in grade 7. Educators conducted the tryouts during regular class time. Items were administered online, using the SurveyMonkey platform, and students were given between one and two hours to complete the trial test.

Psychometric analyses were conducted on these items to identify any items that did not perform as expected. Specifically, the p-value and the item-total correlation (ITC) were used to identify problematic items. For more details on these statistics, see the "Item Statistics" subsection of the "Operational Data Analysis" section. WestEd content experts, WestEd

psychometricians, and PAUSD staff worked in partnership to review items with poor statistics. Items were subsequently removed as needed.

Form Construction

Forms were developed using the blueprint for each course. Items selected for each form demonstrated strong item statistics (i.e., p-value and ITC) from the item tryout analyses. The forms were submitted to PAUSD staff for review and were approved before being used for the cut-score study and the operational test administration.

Test Administration and Scoring

Test Administration

PAUSD students can opt in to participate in the Middle School Math Acceleration Process. All families of rising grades 6, 7, and 8 students were notified of the opportunity to participate. Students may only accelerate one additional year in the PAUSD middle school math pathway. Students may participate more than once if they do not pass the assessments in a previous validation process. Participating in the Math Validation Process is optional.

The validation tests were administered at the three PAUSD middle schools in May 2024, and students took the assessment at their enrolled school. All middle schools administered the validation tests on the same day. The tests took place after school hours, and students had two hours to take the assessments. Students with Individualized Education Plans (IEPs) or 504 Plans designating accommodations for separate settings or extended time were provided access to the accommodations. For most students, their accommodations were extended time, typically 1.5 time (3 hours). For those who had other accommodations, a discussion with families clarified what the accommodation(s) would look like in the validation test setting. A district designee extracted accommodation information for all students with IEPs or 504 Plans who registered for the test. A site designee reviewed the accommodations by student and contacted families to discuss appropriate accommodations for testing.

To ensure that all students were tested under the same standardized testing conditions, test administrators were provided with Directions for Administration (DFA) with test-day protocols and a script with “say boxes” for test administrators to follow throughout the test session. In addition to the test administrator DFA, test administrators projected a slide deck in the test room outlining test-day expectations for students (welcome, purpose of the test, academic honesty statement, bubbling directions, timer, and end-of-test-session procedures).

Multiple-Choice Scoring Procedure

Answer sheets were generated for each student by district personnel and included a bar code with the student’s name and their identification number. This is referred to as a pre-identified (pre-IDed) answer sheet. Each student was provided their pre-IDed answer sheet to bubble in their answer choices. After testing was complete, and to mitigate answer sheet scanning errors, PAUSD instructional leads reviewed answer sheets to ensure that bubbles were marked darkly and that stray marks would not interfere with scanning. Answer sheets were scanned and

digitally sent by each school site to a secure digital folder for electronic scoring by district office personnel.

Free-Response Question Scoring Procedure

Students responded to FRQs in answer books, using paper and pencil. The three middle school math instructional leads collected the answer books (the "Problems to Solve" answer booklets) and gathered them together at one site. As with prior validation tests, the instructional leads were the scorers for the FRQs. Test booklets were equally divided among the scorers to limit any perceived bias. Using the scoring rubrics developed by WestEd, FRQs were scored, and results were marked in the student's test booklet. Scorers then marked each student's pre-IDed FRQ answer sheet with the student's rubric score for each FRQ from their test booklet. If a student did not attempt an answer, it was marked as "N" on the answer sheet, for "no response."

Data and Test Security

Before Test Session Day

Prior to the administration of the validation tests, secure materials included the following: two sets of test booklets (multiple choice and Problems to Solve), MC answer sheets; and FRQ answer sheets. Materials were printed, handled, and stored securely by a limited number of district office personnel. Materials were delivered to each site, to a designated person, by district office personnel. At the sites, testing materials were kept in a secure (locked) location and were only accessible by an instructional lead or site administrator.

During Test Session

Upon check-in at the testing location, each student received their pre-IDed multiple-choice answer sheet. In the secure testing location, test booklets (multiple choice and Problems to Solve) were passed out to students by proctors. During testing, students marked their MC responses on the multiple-choice answer sheet. Students marked their FRQ answers in their Problems to Solve test booklets. If a student finished testing before the end of the testing session, all of the student's testing materials were collected by the proctor before the student was permitted to leave their seat. At the end of the test session, students were directed to close their test booklets and wait for a proctor to collect their test materials before leaving the test session.

End of Test Session

At the end of the test session, secure materials (completed multiple-choice answer sheets, multiple-choice test booklets, and Problems to Solve test booklets) were collected before students left the testing room. The math instructional lead secured all test materials in the test location, then delivered the secure materials to a locked and secure location at their site until the scoring day.

After Test Session

Only math instructional leads (scorers) or site administrators handled the secure materials, with the purpose of either scoring or securing the materials. After scoring, all secure test materials for all sites (completed multiple-choice answer sheets, test booklets with FRQ responses, and completed FRQ answer sheets) were organized and labeled by site and by test. All test materials currently reside in one location within a secure closet where only a site administrator has access.

Operational Data Analysis

This section describes the results and analyses from the operational administration of the validation tests for Foundations, Concepts, and Algebra.

Item Statistics

The item statistics for each test are presented in Tables 4–6. Each table presents the number of students taking that item (N), along with item statistics such as the mean item score, standard deviation (SD), p-value, item-total correlation (ITC), and maximum number of points per item.

An item p-value indicates item difficulty. P-values are bounded between 0 and 1; lower p-values indicate harder items and higher p-values indicate easier items. The p-value for dichotomous items (i.e., items with two score categories, such as correct versus incorrect responses) is the proportion of students responding correctly to the item. All MC items on these validation tests are dichotomous items. The p-value for polytomous items (i.e., items with more than two score categories) is the mean item score divided by the maximum number of points possible on the item.

The ITC is the correlation between the responses on each item and the total score with that item removed (often referred to as the corrected ITC). A high ITC indicates that higher-performing students tend to respond correctly to the item, whereas a low ITC indicates that lower-performing students are more likely to respond correctly. All correlations are bounded between –1 and 1, but for ITCs, we expect the correlation to be positive. A low ITC suggests that the item is not functioning as expected.

For each test, a data review of all of the item statistics was conducted. Items with an ITC below 0.20 were inspected by a team of WestEd psychometricians, WestEd content experts, and PAUSD instructional leads. On each test, one MC item was removed due to a low ITC. Details regarding these items and the data reviews are discussed in the following sections.

Foundations

Table 4 shows the item statistics for the Foundations validation test. There were 293 students who responded to the multiple-choice section of the Foundations test. Due to nonresponses on the FRQs, the N sizes vary for those items. Overall, items functioned as expected. The p-values indicate that items ranged from easy to hard. Almost all items had ITCs above 0.20. The items with ITCs below that threshold were inspected, and only one, *foundations_MC_007*, was

determined not to be functioning as expected. This item was excluded from students' final scores.

Table 4. Item Statistics for Foundations Validation Test

Item ID	N	Mean	SD	P-value	ITC	Max Points
foundations_FRQ_001	276	2.18	1.08	0.73	0.53	3
foundations_FRQ_002	274	1.55	0.67	0.52	0.41	2
foundations_FRQ_003	274	2.39	1.01	0.80	0.48	3
foundations_FRQ_004	275	2.24	1.18	0.75	0.48	3
foundations_FRQ_005	274	2.18	1.04	0.73	0.56	3
foundations_FRQ_006	262	2.23	1.21	0.74	0.61	3
foundations_FRQ_007	255	1.62	1.15	0.54	0.59	3
foundations_FRQ_008	255	1.21	0.91	0.61	0.63	2
foundations_FRQ_009	268	1.14	0.86	0.38	0.52	2
foundations_FRQ_010	275	1.70	1.38	0.57	0.51	3
foundations_FRQ_011	275	2.27	1.11	0.76	0.52	3
foundations_FRQ_012	243	0.83	0.86	0.42	0.48	2
foundations_FRQ_013	263	1.97	1.39	0.66	0.56	3
foundations_FRQ_014	262	0.42	0.49	0.42	0.46	1
foundations_FRQ_015	257	1.45	1.40	0.48	0.67	3
foundations_FRQ_016	277	1.66	0.71	0.55	0.43	2
foundations_FRQ_017	238	0.93	0.93	0.47	0.54	2
foundations_FRQ_018	229	1.35	1.14	0.45	0.55	3
foundations_FRQ_019	262	0.92	0.87	0.46	0.53	2
foundations_FRQ_020	213	1.46	1.69	0.37	0.58	4
foundations_MC_001	293	0.84	0.37	0.84	0.27	1
foundations_MC_002	293	0.69	0.46	0.69	0.45	1
foundations_MC_004	293	0.44	0.50	0.44	0.12	1
foundations_MC_006	293	0.65	0.48	0.65	0.37	1
foundations_MC_007	293	0.59	0.49	0.59	0.04	1
foundations_MC_010	293	0.79	0.41	0.79	0.40	1
foundations_MC_011	293	0.72	0.45	0.72	0.44	1
foundations_MC_012	293	0.81	0.40	0.81	0.30	1

Item ID	N	Mean	SD	P-value	ITC	Max Points
foundations_MC_013	293	0.58	0.49	0.58	0.41	1
foundations_MC_014	293	0.60	0.49	0.60	0.50	1
foundations_MC_015	293	0.71	0.45	0.71	0.37	1
foundations_MC_017	293	0.74	0.44	0.74	0.39	1
foundations_MC_018	293	0.55	0.50	0.55	0.35	1
foundations_MC_019	293	0.49	0.50	0.49	0.58	1
foundations_MC_020	293	0.61	0.49	0.61	0.42	1
foundations_MC_021	293	0.84	0.37	0.84	0.26	1
foundations_MC_023	293	0.71	0.46	0.71	0.49	1
foundations_MC_024	293	0.66	0.47	0.66	0.37	1
foundations_MC_025	293	0.42	0.49	0.42	0.42	1
foundations_MC_028	293	0.84	0.36	0.84	0.14	1
foundations_MC_029	293	0.80	0.40	0.80	0.37	1
foundations_MC_030	293	0.48	0.50	0.48	0.27	1
foundations_MC_031	293	0.80	0.40	0.80	0.43	1
foundations_MC_034	293	0.58	0.49	0.58	0.22	1
foundations_MC_035	293	0.70	0.46	0.70	0.34	1
foundations_MC_039	293	0.62	0.49	0.62	0.44	1
foundations_MC_042	293	0.61	0.49	0.61	0.37	1
foundations_MC_043	293	0.57	0.50	0.57	0.45	1
foundations_MC_045	293	0.62	0.49	0.62	0.48	1
foundations_MC_046	293	0.49	0.50	0.49	0.43	1

Note: The Item ID is a unique item identifier and does not represent the item number in the validation test.

Concepts

Table 5 shows the item statistics for the Concepts validation test. There were 191 students who responded to the multiple-choice section of the Concepts test. Due to nonresponses on the FRQs, the N sizes vary for those items. Overall, items functioned as expected. The p-values indicate that items ranged from easy to hard. Almost all items had ITCs above 0.20. The items with ITCs below that threshold were inspected, and only one, *concepts_MC_007*, was

determined not to be functioning as expected. This item was excluded from students' final scores.

Table 5. Item Statistics for Concepts Validation Test

Item ID	N	Mean	SD	P-value	ITC	Max Points
concepts_FRQ_001	184	1.76	0.97	0.59	0.47	3
concepts_FRQ_002	182	1.74	1.22	0.58	0.55	3
concepts_FRQ_003	164	1.13	1.18	0.38	0.58	3
concepts_FRQ_004	157	0.92	1.27	0.31	0.42	3
concepts_FRQ_005	161	1.71	1.28	0.57	0.62	3
concepts_FRQ_006	143	0.78	0.93	0.26	0.53	2
concepts_FRQ_007	164	1.07	0.88	0.54	0.52	2
concepts_FRQ_008	154	0.82	0.90	0.41	0.58	2
concepts_FRQ_009	144	0.90	0.89	0.45	0.55	2
concepts_FRQ_010	162	1.30	1.31	0.43	0.59	3
concepts_FRQ_011	155	0.72	0.89	0.36	0.60	2
concepts_FRQ_012	135	1.19	0.96	0.40	0.43	2
concepts_FRQ_013	117	0.51	0.76	0.26	0.57	2
concepts_FRQ_014	123	0.78	0.90	0.39	0.64	2
concepts_FRQ_015	153	1.83	1.32	0.46	0.54	4
concepts_FRQ_016	119	0.55	0.73	0.28	0.59	2
concepts_FRQ_017	112	1.24	1.39	0.31	0.69	4
concepts_FRQ_018	121	1.45	1.37	0.36	0.41	3
concepts_FRQ_019	130	2.19	1.31	0.73	0.34	3
concepts_FRQ_020	131	1.15	1.11	0.38	0.51	3
concepts_MC_001	191	0.54	0.50	0.54	0.27	1
concepts_MC_002	191	0.68	0.47	0.68	0.48	1
concepts_MC_003	191	0.21	0.41	0.21	0.20	1
concepts_MC_004	191	0.48	0.50	0.48	0.10	1
concepts_MC_007	191	0.60	0.49	0.60	0.06	1
concepts_MC_009	191	0.48	0.50	0.48	0.26	1
concepts_MC_010	191	0.38	0.49	0.38	0.51	1
concepts_MC_011	191	0.52	0.50	0.52	0.22	1

Item ID	N	Mean	SD	P-value	ITC	Max Points
concepts_MC_012	191	0.42	0.50	0.42	0.30	1
concepts_MC_015	191	0.35	0.48	0.35	0.42	1
concepts_MC_016	191	0.49	0.50	0.49	0.40	1
concepts_MC_017	191	0.64	0.48	0.64	0.24	1
concepts_MC_018	191	0.46	0.50	0.46	0.11	1
concepts_MC_020	191	0.72	0.45	0.72	0.16	1
concepts_MC_023	191	0.48	0.50	0.48	0.45	1
concepts_MC_024	191	0.60	0.49	0.60	0.43	1
concepts_MC_026	191	0.54	0.50	0.54	0.53	1
concepts_MC_027	191	0.66	0.47	0.66	0.27	1
concepts_MC_030	191	0.48	0.50	0.48	0.36	1
concepts_MC_031	191	0.71	0.46	0.71	0.41	1
concepts_MC_032	191	0.44	0.50	0.44	0.42	1
concepts_MC_034	191	0.78	0.42	0.78	0.36	1
concepts_MC_035	191	0.60	0.49	0.60	0.11	1
concepts_MC_036	191	0.49	0.50	0.49	0.52	1
concepts_MC_037	191	0.47	0.50	0.47	0.27	1
concepts_MC_039	191	0.38	0.49	0.38	0.24	1
concepts_MC_041	191	0.32	0.47	0.32	0.38	1
concepts_MC_043	191	0.39	0.49	0.39	0.33	1
concepts_MC_045	191	0.50	0.50	0.50	0.45	1
concepts_MC_046	191	0.40	0.49	0.40	0.41	1

Note: The Item ID is a unique item identifier and does not represent the item number in the validation test.

Algebra

Table 6 shows the item statistics for the Algebra validation test. There were 109 students who responded to the multiple-choice section of the Algebra test. Due to nonresponses on the FRQs, the N sizes vary for those items. Overall, items functioned as expected. The p-values indicate that items ranged from easy to hard. Almost all items had ITCs above 0.20. The items with ITCs below that threshold were inspected, and only one, *algebra_MC_031*, was determined not to be functioning as expected. This item was excluded from students' final scores. One item with a low ITC, *algebra_MC_010*, was determined to have a low ITC due to a statistical artifact (restriction of range) since the item is quite easy. This item was kept on the

test, since the assessment developers purposefully place easy items at the beginning of the test.

Table 6. Item Statistics for Algebra Validation Test

Item ID	N	Mean	SD	P-value	ITC	Max Points
algebra_FRQ_001	83	1.16	0.99	0.39	0.42	3
algebra_FRQ_002	89	1.20	1.17	0.40	0.60	3
algebra_FRQ_003	94	2.03	1.17	0.68	0.69	3
algebra_FRQ_004	76	1.04	0.96	0.52	0.67	2
algebra_FRQ_005	84	1.04	0.91	0.52	0.71	2
algebra_FRQ_006	94	1.86	1.11	0.62	0.50	3
algebra_FRQ_007	95	1.31	1.23	0.44	0.59	3
algebra_FRQ_008	84	1.18	0.98	0.39	0.62	3
algebra_FRQ_009	80	1.05	1.05	0.35	0.48	3
algebra_FRQ_010	73	0.70	1.15	0.23	0.36	3
algebra_FRQ_011	91	0.48	0.50	0.48	0.54	1
algebra_FRQ_012	74	0.59	0.86	0.30	0.45	2
algebra_FRQ_013	83	1.07	0.88	0.54	0.64	2
algebra_FRQ_014	69	1.29	1.14	0.43	0.54	3
algebra_FRQ_015	68	1.06	1.21	0.35	0.50	3
algebra_FRQ_016	83	1.87	1.33	0.62	0.41	3
algebra_FRQ_017	85	1.02	1.41	0.34	0.39	3
algebra_FRQ_018	80	1.18	1.17	0.39	0.41	3
algebra_FRQ_019	89	2.22	1.22	0.74	0.52	3
algebra_FRQ_020	91	2.29	1.25	0.76	0.35	3
algebra_MC_001	109	0.51	0.50	0.51	0.26	1
algebra_MC_005	109	0.76	0.43	0.76	0.40	1
algebra_MC_006	109	0.86	0.35	0.86	0.27	1
algebra_MC_008	109	0.66	0.48	0.66	0.24	1
algebra_MC_009	109	0.87	0.34	0.87	0.46	1
algebra_MC_010	109	0.87	0.34	0.87	-0.05	1
algebra_MC_012	109	0.42	0.50	0.42	0.15	1
algebra_MC_013	109	0.65	0.48	0.65	0.25	1

Item ID	N	Mean	SD	P-value	ITC	Max Points
algebra_MC_014	109	0.76	0.43	0.76	0.31	1
algebra_MC_015	109	0.48	0.50	0.48	0.34	1
algebra_MC_017	109	0.47	0.50	0.47	0.53	1
algebra_MC_018	109	0.55	0.50	0.55	0.55	1
algebra_MC_022	109	0.76	0.43	0.76	0.42	1
algebra_MC_023	109	0.67	0.47	0.67	0.62	1
algebra_MC_024	109	0.83	0.38	0.83	0.22	1
algebra_MC_026	109	0.44	0.50	0.44	0.15	1
algebra_MC_027	109	0.41	0.49	0.41	0.17	1
algebra_MC_029	109	0.47	0.50	0.47	0.36	1
algebra_MC_030	109	0.80	0.40	0.80	0.41	1
algebra_MC_031	109	0.15	0.36	0.15	0.08	1
algebra_MC_033	109	0.63	0.48	0.63	0.34	1
algebra_MC_036	109	0.75	0.43	0.75	0.33	1
algebra_MC_037	109	0.72	0.45	0.72	0.37	1
algebra_MC_039	109	0.61	0.49	0.61	0.38	1
algebra_MC_040	109	0.76	0.43	0.76	0.33	1
algebra_MC_041	109	0.45	0.50	0.45	0.22	1
algebra_MC_042	109	0.58	0.50	0.58	0.22	1
algebra_MC_043	109	0.64	0.48	0.64	0.42	1
algebra_MC_044	109	0.71	0.46	0.71	0.44	1
algebra_MC_046	109	0.49	0.50	0.49	0.30	1

Note: The Item ID is a unique item identifier and does not represent the item number in the validation test.

Reliability

Test reliability is estimated using Cronbach's alpha, a measure of internal consistency that is a function of the number of items, the sum of all the item variances, and the variance of the total scores (Cronbach, 1951). Values range from 0 to 1, with values closer to 1 indicating that the items are closely related and that students score consistently across the items. A value of 0.80 is often used as the threshold for good internal reliability, but higher values are best when consequential decisions are made using test scores. In the case of the validation tests, all three tests have estimated reliability coefficients of 0.92 or higher, as shown in Table 7.

Table 7 also shows the standard error of measurement (SEM) for the student raw scores for each of the validation tests. The SEM is a useful statistic because it helps in understanding the

precision of the test scores and the reliability of the measurement. SEM is calculated as a function of the estimated reliability and the standard deviation of the total scores. As shown in Table 7, the SEMs for the validation tests are 4.97 for Foundations; 4.92 for Concepts; and 4.93 for Algebra.

Table 7. Cronbach's Alpha and Standard Error of Measurement (SEM) for Each Validation Test

	Foundations	Concepts	Algebra
Alpha	0.93	0.92	0.93
SEM	4.97	4.92	4.93

Student Results

Table 8 shows the score distributions for the Foundations, Concepts, and Algebra validation tests. For each test, scores were calculated without the MC items that were determined to be problematic during the item analysis, resulting in a total of 49 items for each test score calculation. Table 8 shows statistics for both raw scores and percent correct for each test.

As shown in Table 8, in the Foundations test, the mean percent correct was 58.49 percent, with a standard deviation of 23.49. In Concepts, the mean percent correct was lower, at 40.27 percent, with a standard deviation of 21.77. In Algebra, the mean percent correct was 46.48 percent, with a standard deviation of 22.19.

Table 8. Student Score Distributions for Raw Scores and Percent Correct on Foundations, Concepts, and Algebra Validation Tests

	Foundations		Concepts		Algebra	
Statistic	**Raw Score**	**Percent Correct**	**Raw Score**	**Percent Correct**	**Raw Score**	**Percent Correct**
N	293	293	191	191	109	109
Mean	47.4	58.49	33.03	40.27	38.53	46.48
SD	18.99	23.44	17.86	21.77	18.44	22.19
Median	48	59	30	37	38	46
Min	5	6	5	6	1	1
Max	81	100	75	91	75	90

The cumulative distributions are shown in the appendix.

Cut-Score Setting Study

Methodology

The goal of the cut-score setting study was to identify for each validation test the score that is associated with successful performance in the course that a student would bypass.

To determine an appropriate cut-score level for each validation test, a modified contrasting-groups approach was used. In the contrasting-groups approach, originally proposed by Livingston and Zieky (1982), the scores of two contrasting groups are compared to determine an appropriate cutoff score. Expert judgments, informed by data on how students who meet external definitions of proficiency perform on the measure, are used to ensure a clear differentiation between those who meet the proficiency standards and those who do not. For the validation tests, each test form was administered to students who were currently on grade in the relevant course (e.g., students currently in an Algebra course took the Algebra test). The external grouping criterion was course grade in the current course (e.g., students who earned an A in Algebra or a B in Algebra the first semester of that course). Then, validation test scores from the on-grade students in the study were grouped by course letter grade (A students and B students) and contrasted. More details regarding the study are provided in the following sections.

Study Sample

As previously described, to determine the appropriate cut score that is associated with successful performance in the course to be bypassed, each validation test was administered to students currently enrolled in the relevant course (i.e., on-grade students). For example, the Foundations validation test was administered to students currently enrolled in the Foundations course.

Nine classrooms were selected at each school for the cut-score sample. Classrooms were selected with the goal of having a range of earned course grades among students. To be included in the cut-score sample, students had to respond to at least half of the FRQs and at least one of the MC items. The numbers of students who met the inclusion criteria in each grade are provided in Table 9 in the following Score Distributions section.

Study Administration

The cut-score study was administered during the school day. Schools had a one-week time period to administer the test. Students were given two hours to complete the tests. Only students with IEPs or 504 Plans with documented extended time were allocated additional time for completion, typically 3 hours (or 1.5 times longer).

Score Distributions

The distributions of scores on each validation test, disaggregated by students' first-semester course grades, are given in Table 9. Both raw scores and percent-correct scores are shown. The items that were removed in the operational scoring process (as described in the previous "Item Statistics" section) were also removed from the cut-score sample scores. In addition to descriptive statistics, density plots were created (see the following Figures 1–3) that show the distributions of students in the on-grade, contrasting-groups cut-score study. The density plots also show the performance on the validation test for the operational administration (i.e., the students who sat for the test that spring to be accelerated to the next grade).

Overall, the averages on each validation test are higher for students with higher grades. For example, on average, students with A grades in the first semester of the course have the highest scores. This indicates that the test is measuring students' course-level math ability as expected. These data, in conjunction with the density plots, informed the cut-score judgments.

Table 9. Raw Score and Percent Correct Distributions for the Cut-Score Sample by Semester 1 Grade on Each Validation Test

Test	Score	Grade	N	M	SD	Median	Min	Max
Foundations	Raw Score	A	49	58.80	12.44	61	26	78
		B	15	45.07	12.14	44	26	70
		C	12	29.50	15.70	26	9	57
		D	1	13.00		13	13	13
	Percent Correct	A	49	72.55	15.34	75	32	96
		B	15	55.53	14.88	54	32	86
		C	12	36.33	19.30	32	11	70
		D	1	16.00		16	16	16
Concepts	Raw Score	A	29	44.28	19.54	43	12	79
		B	11	29.91	15.55	24	13	63
		C	1	9.00		9	9	9
		D	0					
	Percent Correct	A	29	54.00	23.74	52	15	96
		B	11	36.55	19.01	29	16	77
		C	1	11.00		11	11	11
		D	0					
Algebra	Raw Score	A	53	54.09	16.67	54	11	83
		B	10	24.10	8.71	26	8	35
		C	3	28.00	13.45	32	13	39
		D	1	8.00		8	8	8
	Percent Correct	A	53	65.13	20.06	65	13	100
		B	10	29.20	10.36	31.5	10	42
		C	3	34.00	16.09	39	16	47
		D	1	10.00		10	10	10

Figure 1 shows the density plots for students in the cut-score sample who earned an A or a B in the first semester of the Foundations course. Students earning an A are shown on the right side in orange, and are higher scoring, on average. The average percent correct for A students was 72.55 percent. Students earning a B are shown on the left side of the figure, and are lower scoring, on average. The average percent correct for B students was 55.53 percent. The operational density is overlaid across both densities and is shown in gray.

Figure 1. Density Plots of Scores for Students Currently Enrolled in the Foundations Course and Earning an A or a B in the First Semester, Overlaid With the Distribution of the Operational Administration

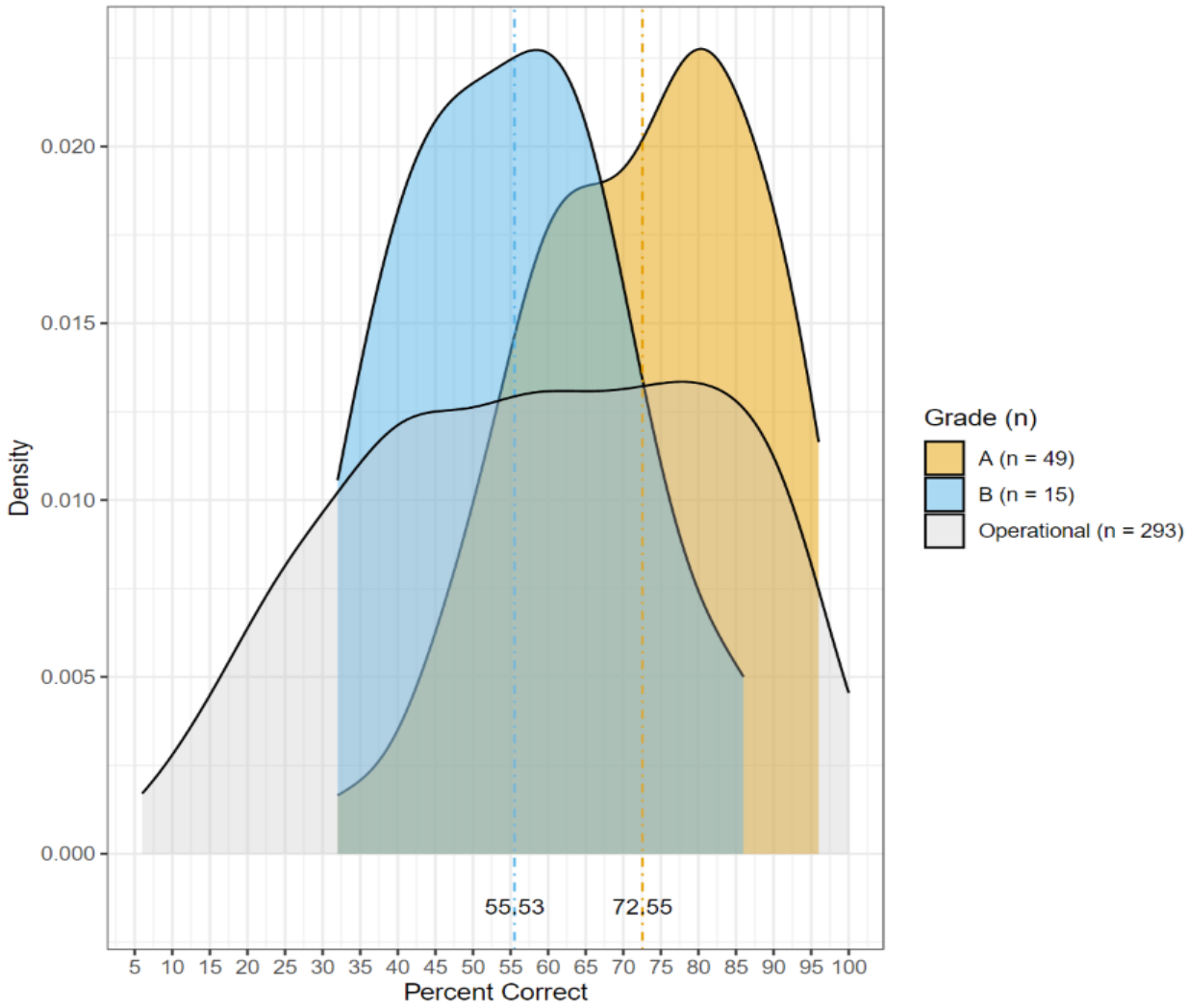


Figure 2 shows the same density plots for Concepts students. The average percent correct for A students was 54 percent, while the average percent correct for B students was 36.64 percent. There is a sizeable overlap between the distributions of the A students and the B students in this validation test. There is also a bimodal distribution, which is most obvious in the A student distribution.

Figure 2. Density Plots of Scores for Students Currently Enrolled in the Concepts Course and Earning an A or a B in the First Semester, Overlaid With the Distribution of the Operational Administration

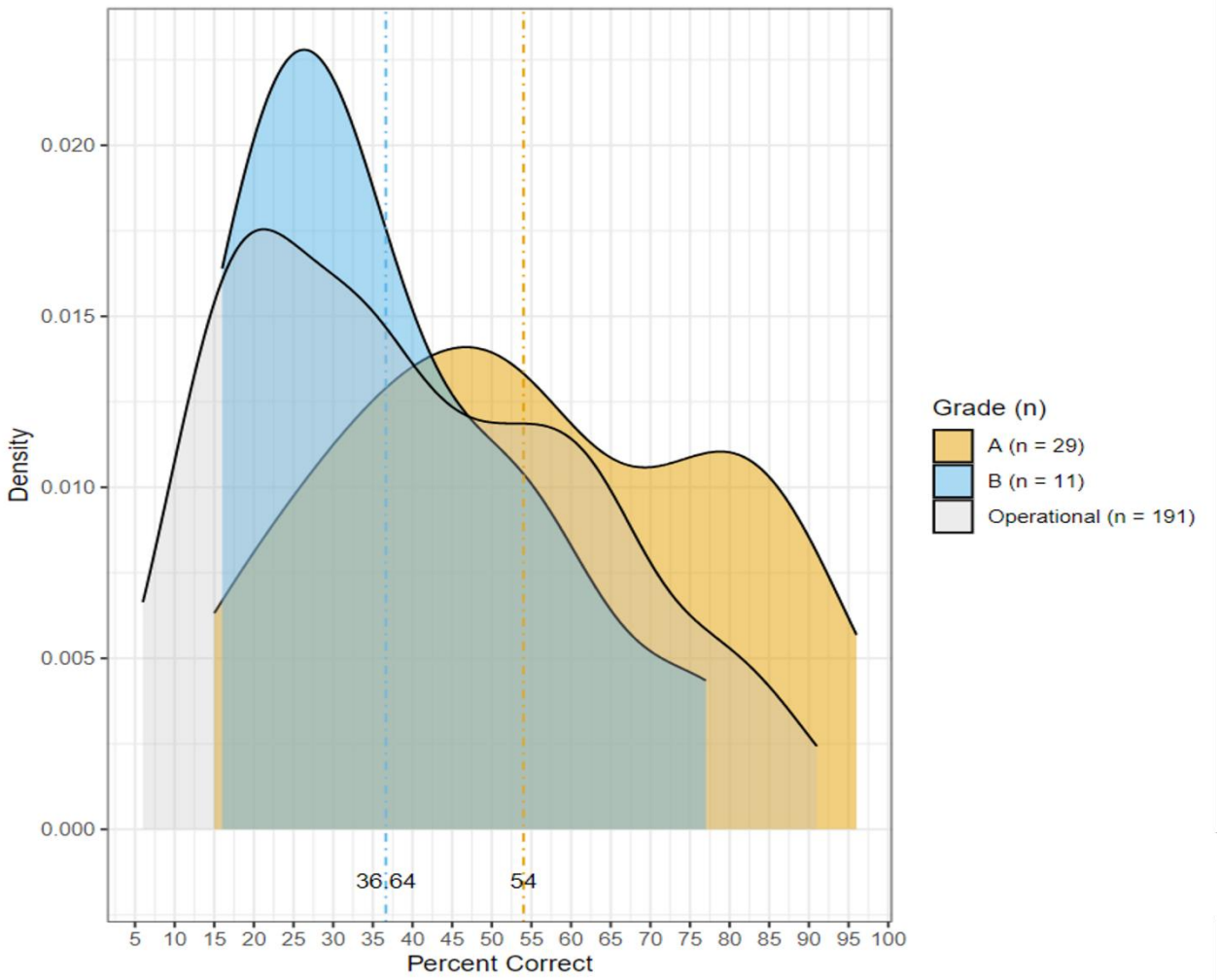
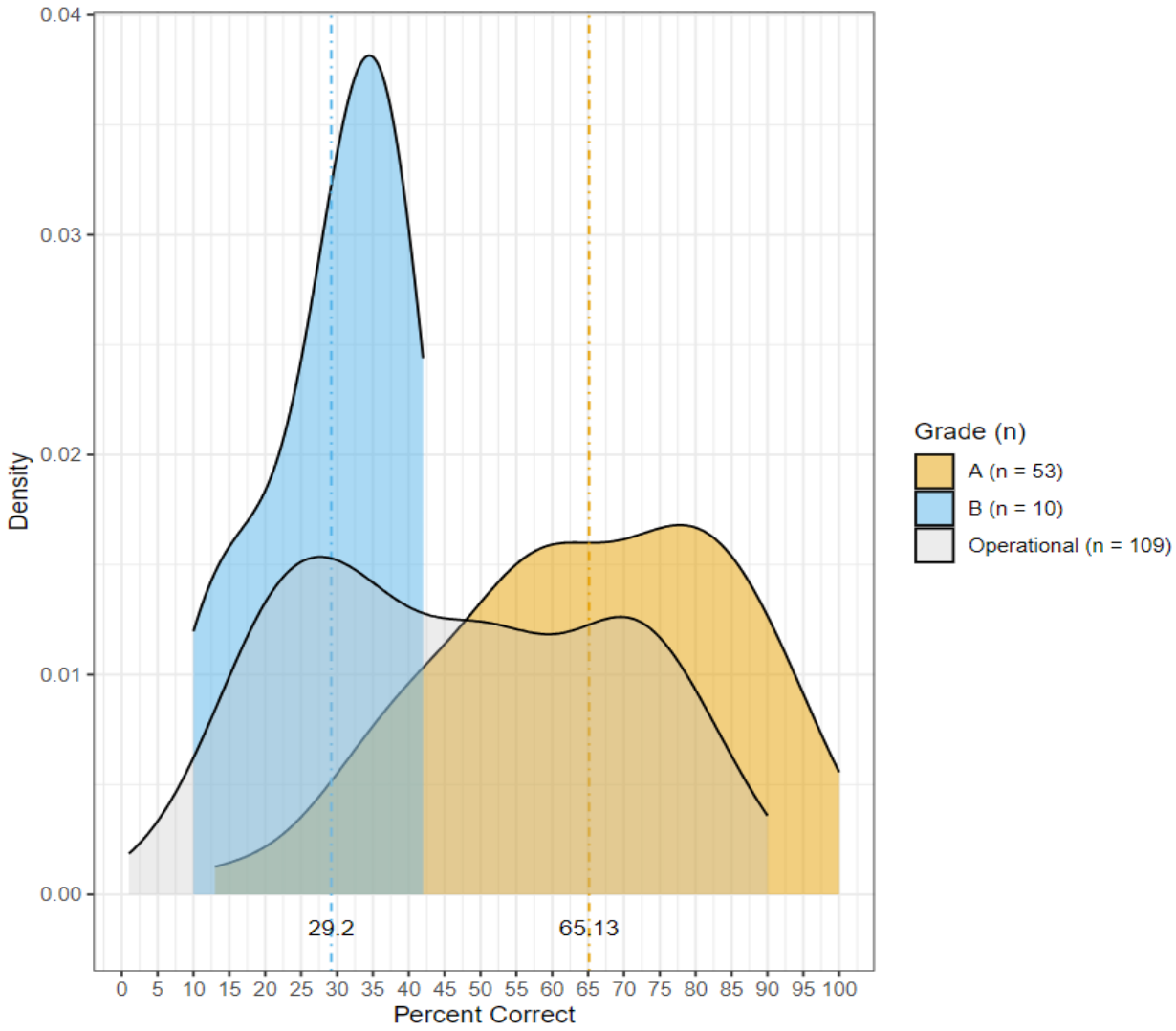


Figure 3 shows the same density plots for Algebra students. The average percents correct were 65.13 percent for A students and 29.2 percent for B students.

Figure 3. Density Plots of Scores for Students Currently Enrolled in the Algebra Course and Earning an A or a B in the First Semester, Overlaid With the Distribution of the Operational Administration



Cut Score and Final Results

The overlapping distributions shown in Figures 1–3 and the cumulative distributions shown in the appendix were presented to PAUSD in order to inform their expert judgments on the final cut-score decisions. Historical pass rates were also compared to the cumulative distributions shown in the appendix. Final decisions were made based on percent correct, such that cut scores represent the minimum percentage of items that a student needs to answer correctly to qualify for acceleration.

Table 10 gives the percentages needed for a student to qualify for bypassing the course. The cut scores for all three exams were determined based on the exam's difficulty and student performance. Statistical analysis indicated that the Concepts validation test was the most challenging for both the cut-score and operational samples, leading to a lower cut score, compared to the Foundations and Algebra tests.

Table 10 also shows the percentages of students who qualified for the course acceleration in each of the three validation tests. In Foundations, 36 percent of students qualified; in Concepts, 23 percent of students qualified; and in Algebra, 20 percent of students qualified.

Table 10. Qualifying Scores, Numbers of Students Participating and Qualifying, and Percentages of Students Qualifying on the Foundations, Concepts, and Algebra Validation Tests

	Foundations	Concepts	Algebra
Qualifying Score	70%	60%	70%
Number of Students Participating	293	191	109
Number of Students Qualifying	106	44	22
Percentage of Students Qualifying	36%	23%	20%

Validity Studies

Two sets of analyses were conducted to gather validity evidence for the use of the established cut scores to identify students who would be successful in the accelerated course.

The first analysis compares student scores on the Smarter Balanced Assessment Consortium (SBAC) math test from spring 2024, the same time period as the validation test administration.

> Results from the first analysis indicate that students who reached the qualifying percent correct perform higher on the grade-appropriate SBAC, on average, compared to students who did not reach the qualifying percent correct.

The second analysis is a predictive validity study that compares first-semester course grades from the 2024/25 academic year.

> Results from the second analysis show that students who bypassed a course based on the validation tests tend to earn strong first-semester course grades in the 2024/25 academic year.

Both analyses are discussed in detail in the following sections.

Comparison to SBAC Math Scale Scores

The first analysis compares students' performance on each validation test to their performance on the grade-appropriate SBAC test. For example, since students taking the Foundations validation test were in grade 5 at the time of the assessment and were candidates for acceleration beyond the grade 6 Foundations course, the mean of the grade 5 math SBAC scores for students who qualified for acceleration is compared to the mean of the grade 5 math SBAC scores for students who did not qualify for acceleration, based on the Foundations validation test.

As shown in Table 11, on the Foundations test, the 127 participants who qualified for acceleration had a mean score of 2699.60, with a standard deviation of 38.92, whereas the 216 participants who did not qualify had a mean score of 2633.91, with a standard deviation of 60.92. A two-sample t-test of mean differences indicates a statistically significant difference on the grade 5 SBAC test for these two groups ($t = 10.9$, $df = 263.95$, $p < 0.001$). Similarly, the Concepts test had 50 participants who qualified for acceleration with a mean grade 6 SBAC

math score of 2729.28 and a standard deviation of 48.43, compared to 165 participants who did not qualify with a mean grade 6 SBAC math score of 2665.65 and a standard deviation of 56.71, again with a statistically significant mean difference (t = 7.26, df = 79.28, p < 0.001). Lastly, the Algebra test had 22 participants who qualified for acceleration with a mean grade 7 SBAC math score of 2765.24 and a standard deviation of 62.28, while the 96 participants who did not qualify had a mean grade 7 SBAC math score of 2715.28 and a standard deviation of 60.43. The mean difference in this group is also statistically significant (t = 3.31, df = 29.88, p = 0.002).

Table 11. Sample Sizes, Means, Standard Deviations, and Two-Sample t-Tests of Mean Difference on SBAC Tests Between Students Who Qualified for Acceleration and Those Who Did Not

Validation Test	SBAC Test Grade	Validation Test Qualification	N	Mean	SD	t	df	p-value
Foundations	Grade 5	Qualified	127	2699.60	38.92			
Foundations	Grade 5	Did Not Qualify	216	2633.91	60.92	10.9	263.95	< 0.001
Concepts	Grade 6	Qualified	50	2729.28	48.43			
Concepts	Grade 6	Did Not Qualify	165	2665.65	56.71	7.26	79.28	< 0.001
Algebra	Grade 7	Qualified	22	2765.24	62.28			
Algebra	Grade 7	Did Not Qualify	96	2715.28	60.43	3.31	29.88	0.002

Figures 4, 5, and 6 show the density plots for these distributions for the Foundations, Concepts, and Algebra assessments, respectively. In each of the plots, the distribution on the right, in orange, represents the distribution of students who qualified for acceleration, and the distribution on the left, in blue, represents the distribution of students who did not qualify for acceleration. The dashed lines through each distribution show the mean scores.

Figure 4. Density Plot of Grade 5 SBAC Math Scale Scores by Acceleration Status for Students Taking the Foundations Test

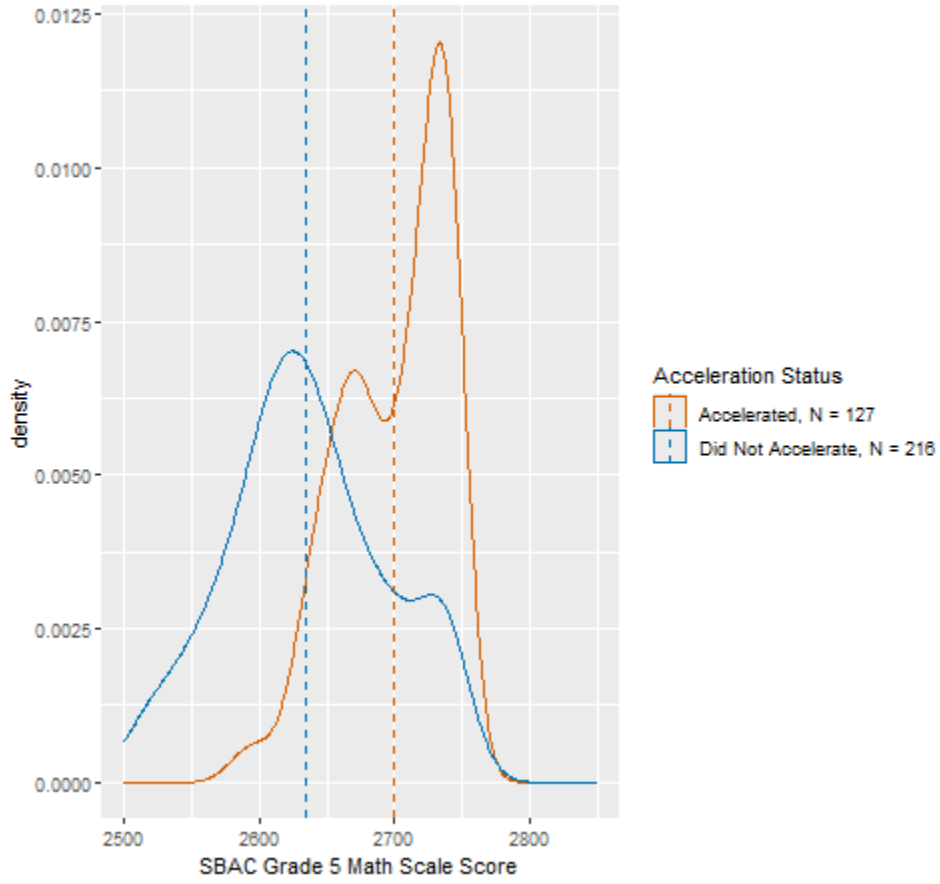


Figure 5. Density Plot of Grade 6 SBAC Math Scale Scores by Acceleration Status for Students Taking the Concepts Test

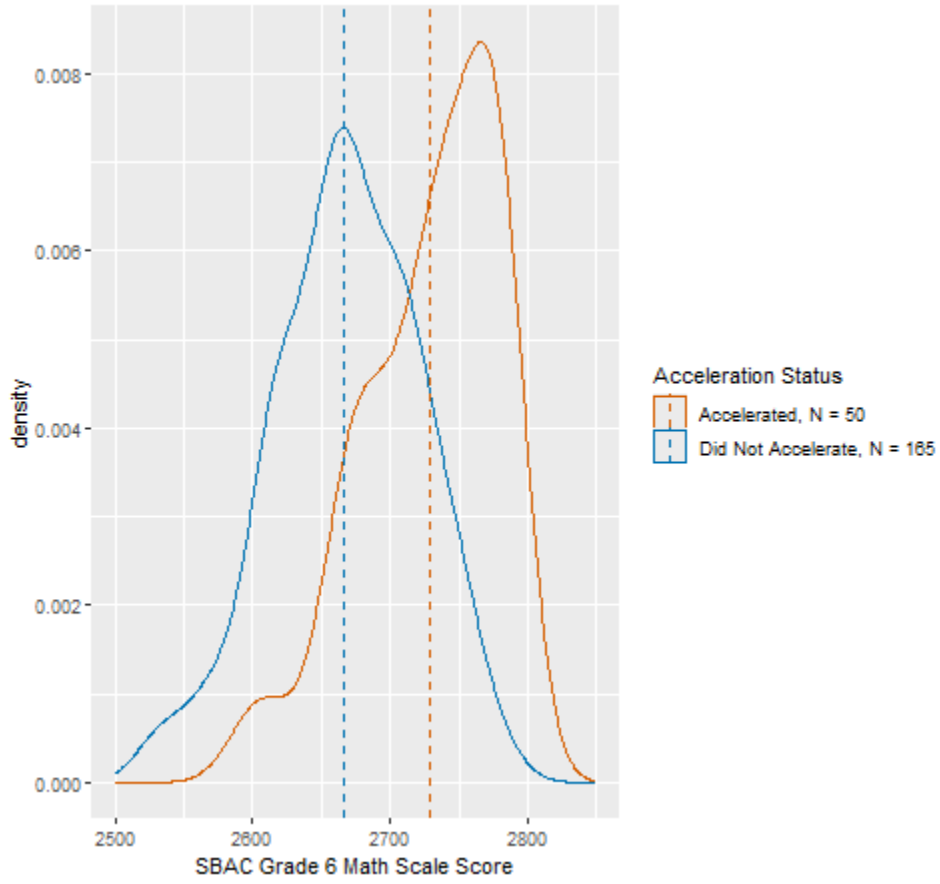
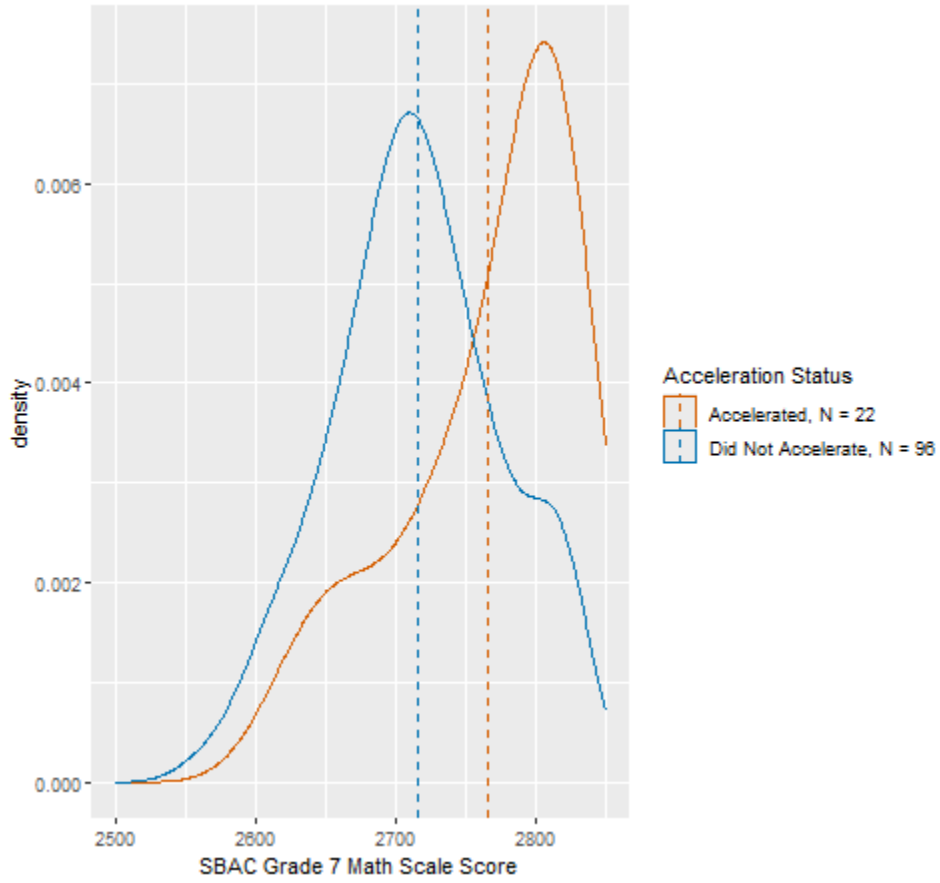


Figure 6. Density Plot of Grade 7 SBAC Math Scale Scores by Acceleration Status for Students Taking the Algebra Test



In addition to comparing the distributions by qualification status, examining the relationship between SBAC math scale scores and percents correct on the validation tests offers evidence regarding the appropriateness of the established cut scores. The purpose of analyzing the relationships between validation test performance and SBAC performance is to evaluate whether students who demonstrate proficiency on above-grade content and score above the validation test cut scores also demonstrate strong performance on grade-level content.

Figure 7 shows a scatter plot of the grade 5 SBAC math scale scores and the validation test percents correct. The lower dotted horizontal line, in blue, shows the California state mean for the grade 5 SBAC scores, while the upper dotted horizontal line, in green, shows the PAUSD mean score. Overall, PAUSD students performed well above the state mean on the grade 5 SBAC. The vertical line, in purple, shows the set cut score on the validation test. Students to the right of the line qualified for acceleration. This plot indicates a linear relationship between the content assessed on the grade 5 SBAC and the grade 6 content assessed on the Foundations

validation test, with the former measuring grade-level standards and the latter assessing readiness for advanced content.

Figure 8 shows the relationship between Concepts validation test scores and grade 6 SBAC math scale scores, and Figure 9 shows the relationship between Algebra validation test scores and grade 7 SBAC math scale scores. Both plots have the same vertical and horizontal lines as the Foundations plot. In Concepts, students who demonstrated strong performance on the validation test also demonstrated strong performance on the grade 6 SBAC math assessment. Most students taking the Concepts test scored above the PAUSD mean. In Algebra, students to the right of the cut-score line demonstrated high performance on the grade 7 SBAC test. These plots show linear relationships between the content assessed on the grade-level SBAC tests and the advanced content assessed on the respective validation tests, with the SBAC tests measuring current grade-level standards and the validation tests assessing readiness for accelerated coursework.

Figure 7. Percent Correct on Validation Test versus Grade 5 SBAC Math Scale Score for Students Taking the Foundations Test

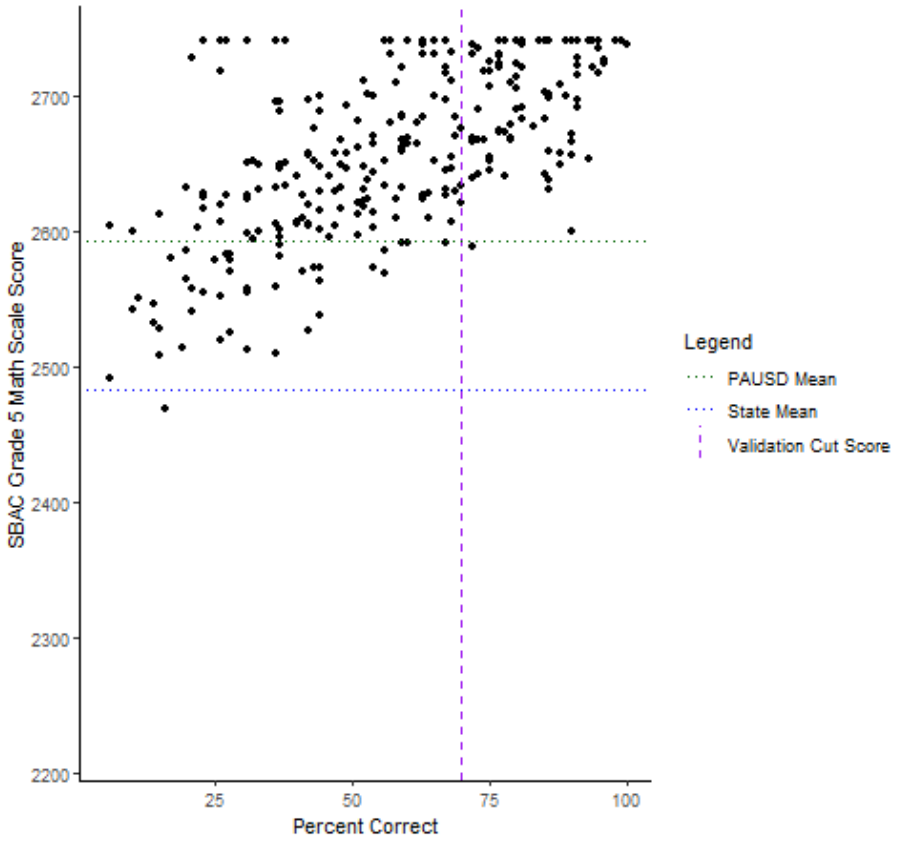


Figure 8. Percent Correct on Validation Test versus Grade 6 SBAC Math Scale Score for Students Taking the Concepts Test

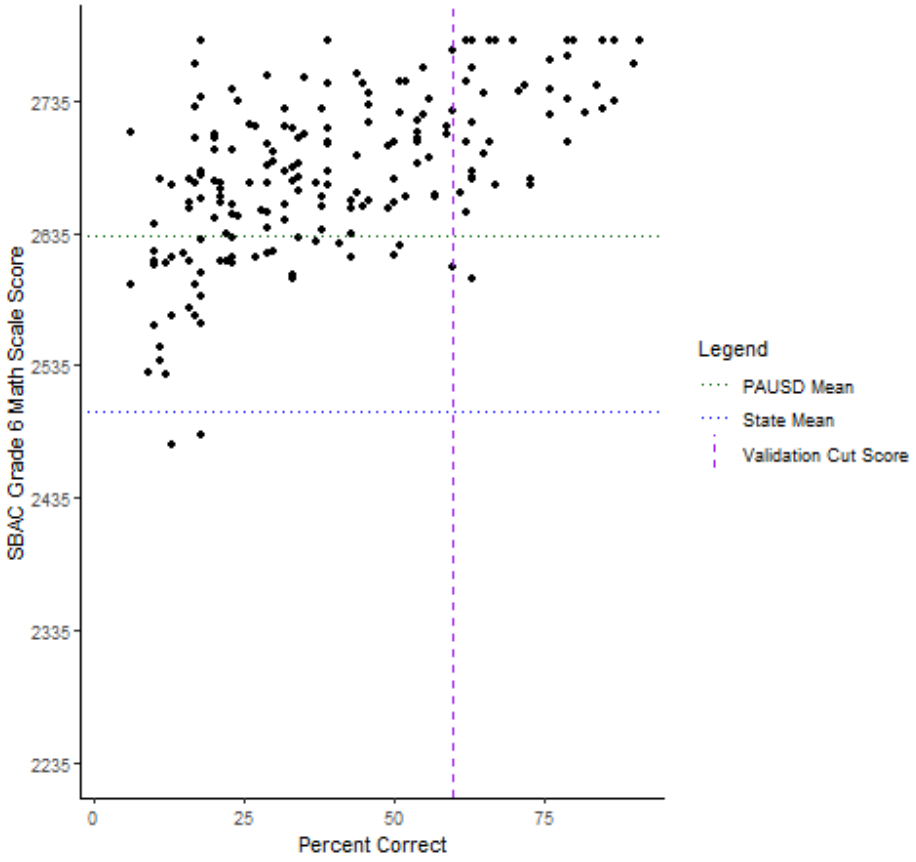
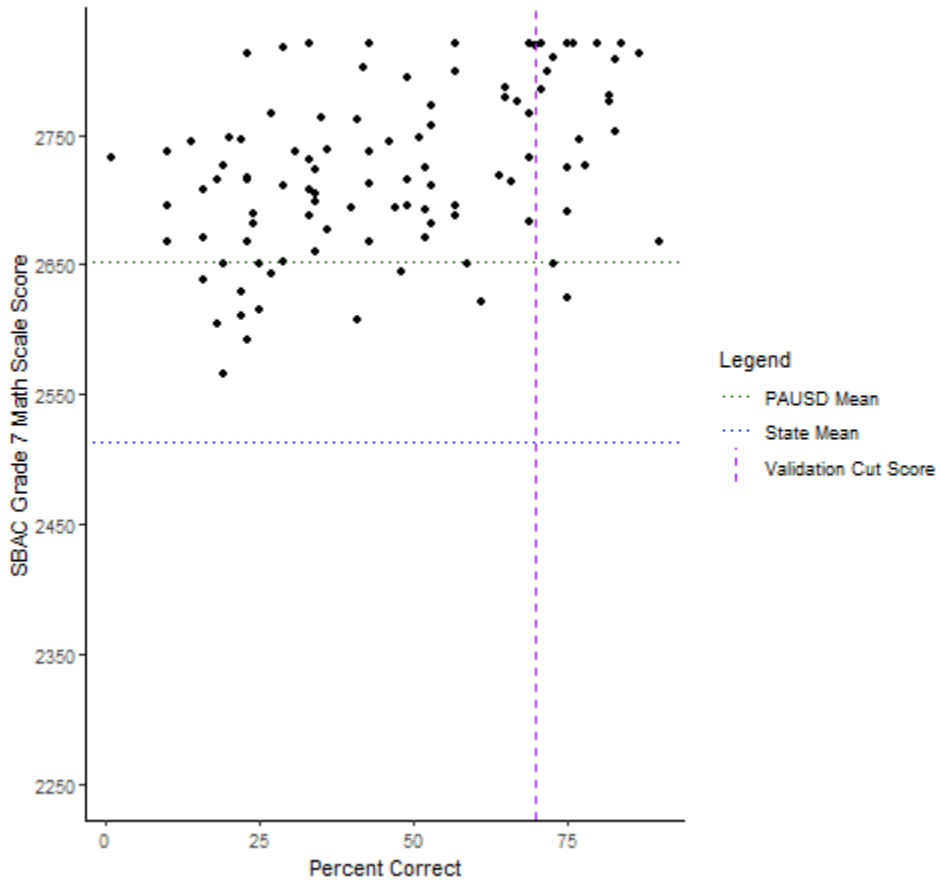


Figure 9. Percent Correct on Validation Test versus Grade 7 SBAC Math Scale Score for Students Taking the Algebra Test



Comparison of Semester Course Grades

The second set of analyses compares the first-semester course grades in the 2024/25 academic year. The purpose of this analysis was to evaluate whether students who accelerate after meeting qualification criteria on the validation tests demonstrate achievement levels comparable to or exceeding those of students who completed the prerequisite course sequence.

Figure 10 shows the proportions of students per letter grade, separated by acceleration status. Students who accelerated into the Concepts course bypassed the Foundations course, whereas students who did not accelerate into Concepts completed the prerequisite Foundations course.

Students who bypassed Foundations exhibited strong academic performance in the Concepts course.

Figure 10. Proportion of Students per Letter Grade, by Acceleration Status, for Students Enrolled in Concepts Course

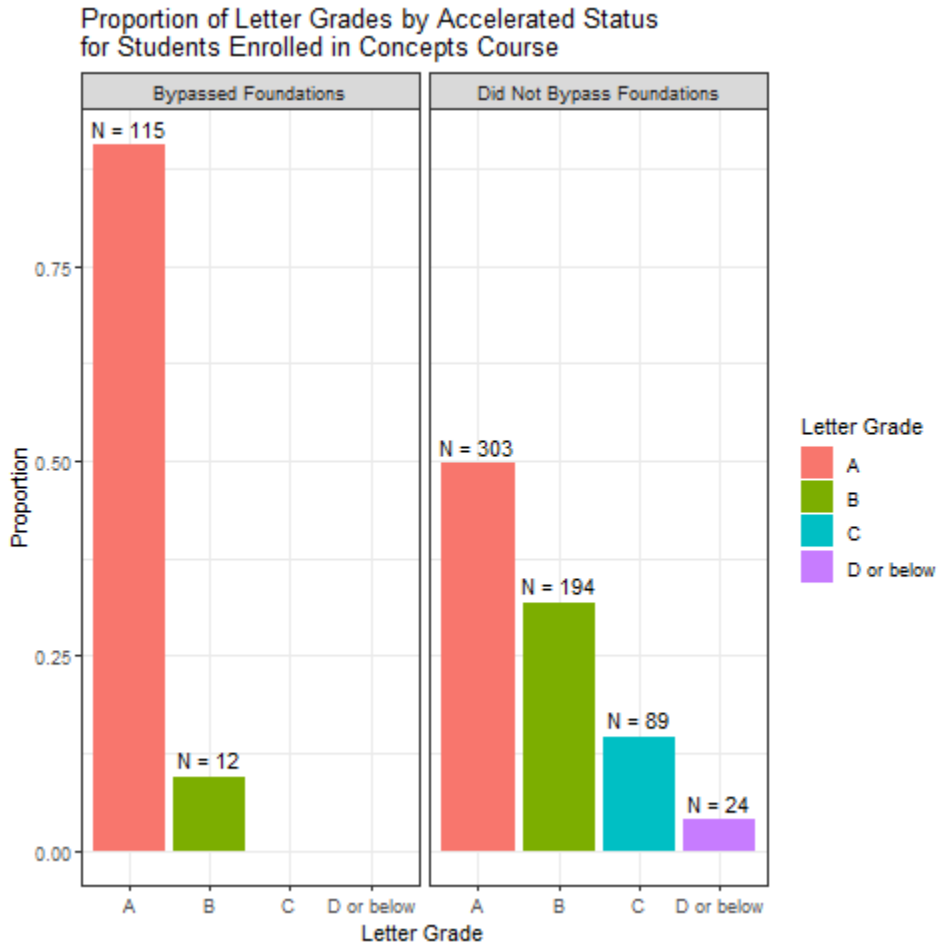
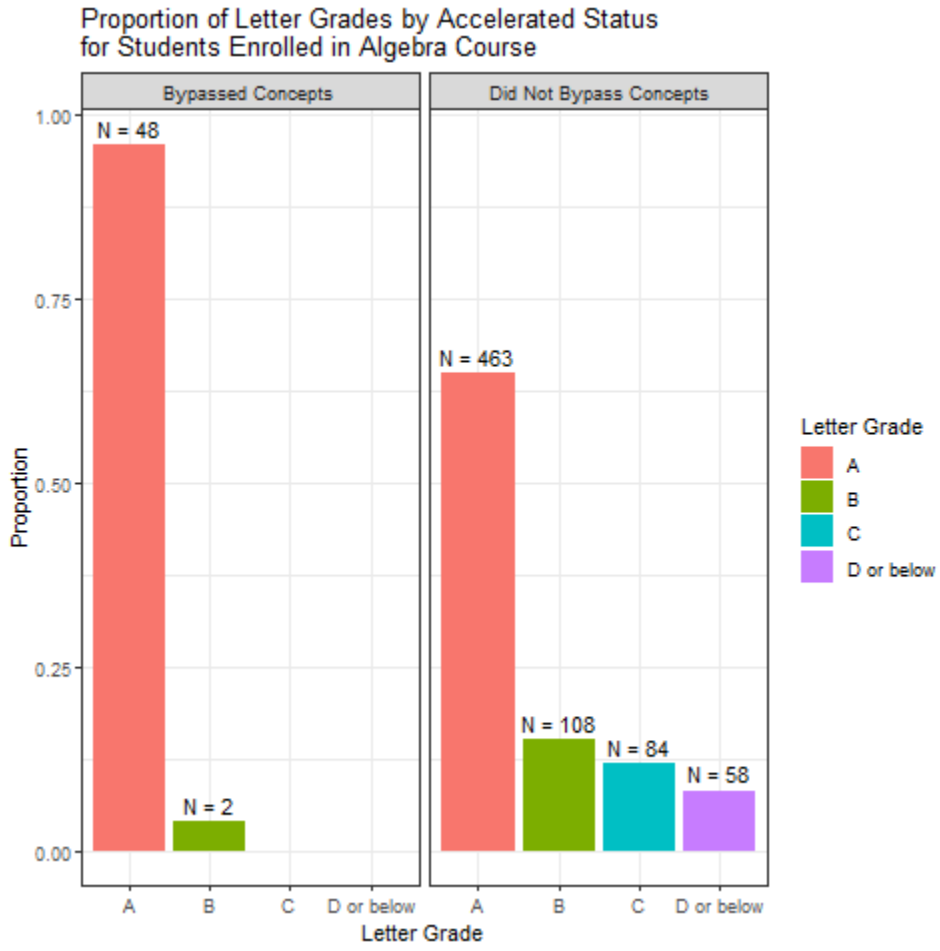


Figure 11 shows the same data for students enrolled in Algebra. Students who accelerated into the Algebra course bypassed the Concepts course, whereas students who did not accelerate into Algebra completed the prerequisite Concepts course. Students who bypassed Concepts exhibited strong academic performance in the Algebra course.

Figure 11. Proportion of Students per Letter Grade, by Accelerated Status, for Students Enrolled in Algebra Course



Lastly, Figure 12 shows the proportions of Geometry students per letter grade, separated by acceleration status. Students who accelerated into the Geometry course bypassed the Algebra course, whereas students who did not accelerate into Geometry completed the prerequisite Algebra course. It should be noted that for students to be enrolled in Geometry in middle school, those who completed the prerequisite Algebra course would have bypassed either the Foundations course or the Concepts course in their academic pathway. The data suggest that students who bypassed Algebra demonstrate generally comparable academic performance to

their peers, with similar proportions of students earning a B or better across both groups. However, students who completed Algebra prior to Geometry show a higher proportion of A grades.

Figure 12. Proportion of Students per Letter Grade, by Accelerated Status, for Students Enrolled in Geometry Course

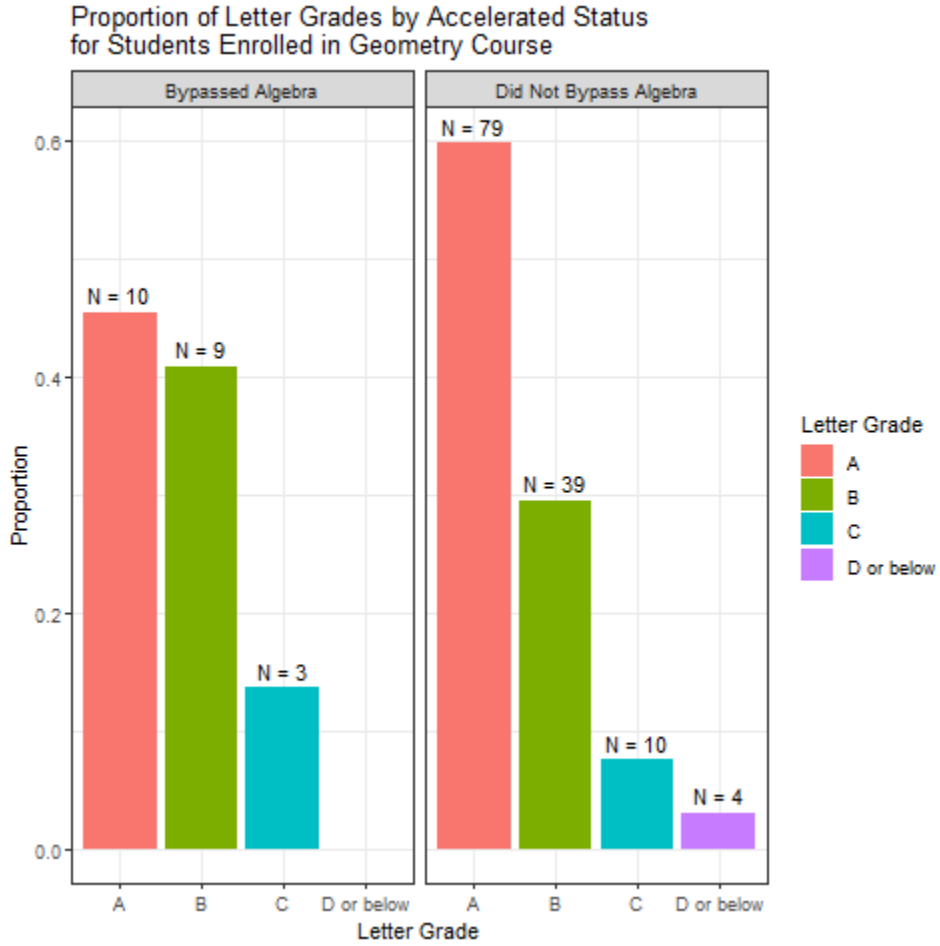


Table 12 presents the distribution of grades (A through D or below) for students in each course, categorized by their acceleration status. The table shows three student groups: those who qualified for acceleration through a validation test in 2024, those who took a validation test in 2024 but did not qualify, and those who followed the standard pathway without testing in 2024. For Concepts and Algebra courses, all three pathways are represented, while Geometry shows only students who qualified via the Algebra validation test and those who did not test, (since there is no higher validation test for this level). This table provides a further disaggregation of the numerical data underlying the grade distributions depicted in Figures 10–12.

Table 12. Counts of Students Within Each Course by Acceleration Status for Each Validation Test

Current Course	Acceleration Status	Test Taken	A	B	C	D or below
Concepts	Qualified	Foundations	115	12	0	0
Concepts	Did Not Qualify	Concepts	108	49	8	0
Concepts	Did Not Test		195	145	81	24
Algebra	Qualified	Concepts	48	2	0	0
Algebra	Did Not Qualify	Algebra	88	8	0	0
Algebra	Did Not Test		375	100	84	58
Geometry	Qualified	Algebra	10	9	3	0
Geometry	Did Not Test		79	39	10	4

Summary

Overall, the analyses in this section provide validity evidence supporting the use of the established cut scores for identifying students who will be successful in accelerated mathematics courses. The first analysis demonstrates a positive relationship between performance on the validation tests and achievement on grade-level SBAC mathematics assessments, though it is important to note that these instruments measure different constructs. The second analysis provides evidence that students who meet qualification criteria on the validation tests generally demonstrate strong academic performance in their subsequent accelerated courses.

As with all educational research, there are inherent constraints in the available data and research design. Student performance is measured across different assessments and courses, each with distinct objectives and contexts. Directions for future research could include longitudinal analyses as the cohort progresses through the mathematics sequence. The current analyses demonstrate positive associations between validation test performance and both standardized assessment results and classroom achievement, providing evidence for the validation test cut scores and their use in the PAUSD mathematics program.

References

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

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, *16*(3), 297–334.

Livingston, S. A., & Zieky, M. J. (1982). *Passing scores: A manual for setting standards of performance on educational and occupational tests*. Educational Testing Service.

Appendix

Table A1. Cumulative Frequencies for Foundations Validation Test

Percent Correct	Frequency	Cumulative Frequency	Cumulative Proportion	Reverse Cumulative Frequency	Reverse Cumulative Proportion
6	2	2	0.68	293	100
10	2	4	1.37	291	99.32
11	1	5	1.71	289	98.63
14	2	7	2.39	288	98.29
15	3	10	3.41	286	97.61
16	1	11	3.75	283	96.59
17	1	12	4.1	282	96.25
19	1	13	4.44	281	95.9
20	3	16	5.46	280	95.56
21	3	19	6.48	277	94.54
22	1	20	6.83	274	93.52
23	5	25	8.53	273	93.17
25	1	26	8.87	268	91.47
26	6	32	10.92	267	91.13
27	3	35	11.95	261	89.08
28	4	39	13.31	258	88.05
31	8	47	16.04	254	86.69
32	2	49	16.72	246	83.96
33	3	52	17.75	244	83.28
36	6	58	19.8	241	82.25
37	8	66	22.53	235	80.2
38	4	70	23.89	227	77.47
40	3	73	24.91	223	76.11
41	4	77	26.28	220	75.09
42	7	84	28.67	216	73.72
43	3	87	29.69	209	71.33
44	10	97	33.11	206	70.31

Percent Correct	Frequency	Cumulative Frequency	Cumulative Proportion	Reverse Cumulative Frequency	Reverse Cumulative Proportion
46	2	99	33.79	196	66.89
47	3	102	34.81	194	66.21
48	4	106	36.18	191	65.19
49	3	109	37.2	187	63.82
51	5	114	38.91	184	62.8
52	4	118	40.27	179	61.09
53	4	122	41.64	175	59.73
54	7	129	44.03	171	58.36
56	7	136	46.42	164	55.97
57	3	139	47.44	157	53.58
58	3	142	48.46	154	52.56
59	8	150	51.19	151	51.54
60	4	154	52.56	143	48.81
62	2	156	53.24	139	47.44
63	6	162	55.29	137	46.76
64	3	165	56.31	131	44.71
65	5	170	58.02	128	43.69
67	9	179	61.09	123	41.98
68	5	184	62.8	114	38.91
69	3	187	63.82	109	37.2
70	3	190	64.85	106	36.18
72	6	196	66.89	103	35.15
73	4	200	68.26	97	33.11
74	2	202	68.94	93	31.74
75	7	209	71.33	91	31.06
77	7	216	73.72	84	28.67
78	4	220	75.09	77	26.28
79	5	225	76.79	73	24.91
80	5	230	78.5	68	23.21
81	6	236	80.55	63	21.5
83	2	238	81.23	57	19.45
84	3	241	82.25	55	18.77
85	7	248	84.64	52	17.75

Percent Correct	Frequency	Cumulative Frequency	Cumulative Proportion	Reverse Cumulative Frequency	Reverse Cumulative Proportion
86	6	254	86.69	45	15.36
88	3	257	87.71	39	13.31
89	5	262	89.42	36	12.29
90	6	268	91.47	31	10.58
91	7	275	93.86	25	8.53
93	4	279	95.22	18	6.14
94	3	282	96.25	14	4.78
95	4	286	97.61	11	3.75
96	3	289	98.63	7	2.39
98	1	290	98.98	4	1.37
99	2	292	99.66	3	1.02
100	1	293	100	1	0.34

Table A2. Cumulative Frequencies for Concepts Validation Test

Percent Correct	Frequency	Cumulative Frequency	Cumulative Proportion	Reverse Cumulative Frequency	Reverse Cumulative Proportion
6	3	3	1.57	191	100
9	1	4	2.09	188	98.43
10	5	9	4.71	187	97.91
11	3	12	6.28	182	95.29
12	2	14	7.33	179	93.72
13	4	18	9.42	177	92.67
15	1	19	9.95	173	90.58
16	4	23	12.04	172	90.05
17	7	30	15.71	168	87.96
18	10	40	20.94	161	84.29
20	5	45	23.56	151	79.06
21	6	51	26.7	146	76.44
22	2	53	27.75	140	73.3
23	6	59	30.89	138	72.25
24	2	61	31.94	132	69.11
26	2	63	32.98	130	68.06
27	2	65	34.03	128	67.02
28	1	66	34.55	126	65.97
29	6	72	37.7	125	65.45
30	4	76	39.79	119	62.3
32	5	81	42.41	115	60.21
33	5	86	45.03	110	57.59
34	5	91	47.64	105	54.97
35	4	95	49.74	100	52.36
37	2	97	50.79	96	50.26
38	4	101	52.88	94	49.21
39	6	107	56.02	90	47.12
41	1	108	56.54	84	43.98
43	4	112	58.64	83	43.46
44	3	115	60.21	79	41.36
45	2	117	61.26	76	39.79

Percent Correct	Frequency	Cumulative Frequency	Cumulative Proportion	Reverse Cumulative Frequency	Reverse Cumulative Proportion
46	4	121	63.35	74	38.74
49	3	124	64.92	70	36.65
50	4	128	67.02	67	35.08
51	3	131	68.59	63	32.98
52	2	133	69.63	60	31.41
54	5	138	72.25	58	30.37
55	2	140	73.3	53	27.75
56	2	142	74.35	51	26.7
57	2	144	75.39	49	25.65
59	3	147	76.96	47	24.61
60	3	150	78.53	44	23.04
61	1	151	79.06	41	21.47
62	3	154	80.63	40	20.94
63	8	162	84.82	37	19.37
65	2	164	85.86	29	15.18
66	3	167	87.43	27	14.14
67	2	169	88.48	24	12.57
70	1	170	89.01	22	11.52
71	1	171	89.53	21	10.99
72	1	172	90.05	20	10.47
73	2	174	91.1	19	9.95
76	3	177	92.67	17	8.9
79	5	182	95.29	14	7.33
80	1	183	95.81	9	4.71
82	1	184	96.34	8	4.19
84	1	185	96.86	7	3.66
85	1	186	97.38	6	3.14
87	3	189	98.95	5	2.62
90	1	190	99.48	2	1.05
91	1	191	100	1	0.52

Table A3. Cumulative Frequencies for Algebra Validation Test

Percent Correct	Frequency	Cumulative Frequency	Cumulative Proportion	Reverse Cumulative Frequency	Reverse Cumulative Proportion
1	1	1	0.92	109	100
10	3	4	3.67	108	99.08
14	1	5	4.59	105	96.33
16	3	8	7.34	104	95.41
18	2	10	9.17	101	92.66
19	3	13	11.93	99	90.83
20	1	14	12.84	96	88.07
22	3	17	15.6	95	87.16
23	5	22	20.18	92	84.4
24	2	24	22.02	87	79.82
25	2	26	23.85	85	77.98
27	3	29	26.61	83	76.15
29	3	32	29.36	80	73.39
31	1	33	30.28	77	70.64
33	5	38	34.86	76	69.72
34	5	43	39.45	71	65.14
35	1	44	40.37	66	60.55
36	2	46	42.2	65	59.63
40	1	47	43.12	63	57.8
41	2	49	44.95	62	56.88
42	1	50	45.87	60	55.05
43	4	54	49.54	59	54.13
46	1	55	50.46	55	50.46
47	1	56	51.38	54	49.54
48	1	57	52.29	53	48.62
49	3	60	55.05	52	47.71
51	2	62	56.88	49	44.95
52	3	65	59.63	47	43.12
53	4	69	63.3	44	40.37
57	5	74	67.89	40	36.7
59	1	75	68.81	35	32.11

Percent Correct	Frequency	Cumulative Frequency	Cumulative Proportion	Reverse Cumulative Frequency	Reverse Cumulative Proportion
61	1	76	69.72	34	31.19
64	1	77	70.64	33	30.28
65	2	79	72.48	32	29.36
66	2	81	74.31	30	27.52
67	1	82	75.23	28	25.69
69	5	87	79.82	27	24.77
70	1	88	80.73	22	20.18
71	3	91	83.49	21	19.27
72	1	92	84.4	18	16.51
73	2	94	86.24	17	15.6
75	4	98	89.91	15	13.76
76	1	99	90.83	11	10.09
77	1	100	91.74	10	9.17
78	1	101	92.66	9	8.26
80	1	102	93.58	8	7.34
82	2	104	95.41	7	6.42
83	2	106	97.25	5	4.59
84	1	107	98.17	3	2.75
87	1	108	99.08	2	1.83
90	1	109	100	1	0.92


WestEd
WestEd.org