RENAISSANCE®

Star Assessments™ for Math—High School Technical Manual

RENAISSANCE
Star
Math®

Renaissance Learning
PO Box 8036
Wisconsin Rapids, WI 54495-8036
Telephone: (800) 338-4204
(715) 424-3636

Outside the US: 1.715.424.3636
Fax: (715) 424-4242
Email (general questions): answers@renaissance.com
Email (technical questions): support@renaissance.com
Email (international support): worldsupport@renaissance.com
Website: www.renaissance.com

Copyright Notice

Copyright © 2021 Renaissance Learning, Inc. All Rights Reserved.

This publication is protected by US and international copyright laws. It is unlawful to duplicate or reproduce any copyrighted material without authorization from the copyright holder. This document may be reproduced only by staff members in schools that have a license for the Star Math software. For more information, contact Renaissance Learning, Inc. at the address shown above.

All logos, designs, and brand names for Renaissance's products and services, including, but not limited to, Accelerated Math, Accelerated Reader, Accelerated Reader 360, AccelScan, English in a Flash, MathFacts in a Flash, myON, myON Reader, myON News, Renaissance, Renaissance Flow 360, Renaissance Growth Platform, Renaissance Growth Alliance, Renaissance Learning, Renaissance-U, Renaissance Smart Start, Star, Star 360, Star Custom, Star Early Literacy, Star Early Literacy Spanish, Star Math, Star Math Spanish, Star Reading, Star Reading Spanish and Star Spanish, are trademarks of Renaissance. All other product and company names should be considered the property of their respective companies and organizations.

Macintosh is a trademark of Apple Inc., registered in the U.S. and other countries.

3/2021 SMHS

Contents

Introduction

Star Math High School: Screening Assessments

Since the 2013–2014 school year, two Star Math high school tests—Algebra and Geometry—have been available for use in assessing the mathematical abilities of students in grades 7–12. Both the Algebra and the Geometry tests are 34-item standards-based adaptive assessments, aligned to state and national curriculum standards, that take an average of less than 30 minutes. The high school tests provide immediate feedback to teachers and administrators on each student's mathematical ability in Algebra or Geometry.

Star Math High School Purpose

As periodic screening assessments, Star Math High School tests serve three purposes. First, they provide educators with quick and accurate estimates of students' instructional math levels. Second, they assess math levels relative to norms from a national sample. Third, they provide the means for tracking growth in a consistent manner longitudinally for all students. This is especially helpful to school- and district-level administrators.

While the Star Math High School tests provide accurate normed data like traditional norm-referenced tests, they are not intended to be used as "high-stakes" tests. Generally, states use high-stakes assessments to document growth, adequate yearly progress, and mastery of state standards. These high-stakes tests are also used to report end-of-period performance to parents and administrators or to determine eligibility for promotion or placement. Star Math High School tests are not intended for these purposes. Rather, when high correlations between the Star Math High School tests and high-stakes instruments exist, classroom teachers can use Star Math High School scores to fine-tune instruction while there is still time to improve performance before the regular test cycle. At the same time, school- and district-level administrators can use Star Math High School scores to predict performance on high-stakes tests when there is an existing linking study relating Star Math High School to those tests. Furthermore, Star Math High School results can easily be disaggregated to identify and address the needs of various groups of students.

The Star Math High School tests' repeatability and flexible administration provide specific advantages for everyone responsible for the education process:

- For students, each Star Math High School test's software provides a challenging, interactive, and brief test that builds confidence in their math ability.
- For teachers, the Star Math High School tests facilitate individualized instruction by identifying students who need remediation or enrichment most.
- For principals, the Star Math High School software provides regular, accurate reports on performance at the class, grade, and building level.
- For district administrators and assessment specialists, it provides a wealth of reliable and timely data on math growth at each school and districtwide. It also provides a valid basis for comparing data across schools, grades, and special student populations.

This manual documents the suitability of Star Math High School computer-adaptive testing for these purposes and demonstrates quantitatively how well these innovative instruments in math assessment perform.

Design of Star Math High School

First Generation of Star Math High School Assessments

The introduction of the current version of Star Math High School tests in 2013 marked the first generation of Star Math High School assessments. Star Math High School's original item bank contained 1,700+ items spanning more than 185 objectives.

The Star Math High School tests are standards-based tests. The organization of the content in Star Math High School tests differs from that of the general Star Math test that spans grade K–12. Star Math High School tests' content organization reflects current thinking, as embodied in many different sets of national and local curriculum standards. Two tests, developed to match high school standards for students taking Algebra 1 and 2 as well as Geometry courses are included in Star Math High School: Algebra and Geometry. For the Algebra test, 196 skills are grouped into 24 blueprint skill sets. For the Geometry test, 95 skills are grouped into 13 blueprint skill sets. Some of the skills included in the test blueprint are identified as prerequisite skills normally taught at earlier grades than High School Algebra and Geometry, and items measuring the prerequisite skills may be administered as part of the tests.

These prerequisite skills items tend to be lower in calibrated difficulty than the high school algebra and geometry items, and are included for the purpose of assessing students who are performing below high school level.

The Star Math High School tests are administered as separate 34-item tests, one of Algebra and one of Geometry. For High School Algebra, the test blueprint prescribes allowable ranges of the number of items to be administered from each of 24 skill sets. For each Algebra skill set, prerequisite skills (if any) are also listed; any student's test may include, prerequisite skills items in place of High School Algebra skill set items, if the student's performance indicates that items of lower difficulty than the High School Algebra items may be appropriate. Within each skill set, a maximum number of items from any single skill is prescribed; the intent of this is to ensure that each test spans a range of content, rather than a concentration of multiple items measuring the same skill.

For High School Geometry, the test blueprint prescribes allowable ranges of the number of items to be administered from each of 13 skill sets. For each Geometry skill set, prerequisite skills (if any) are also listed; any student's test may include prerequisite skills items in place of High School Geometry skill set items, if the student's performance is commensurate for items of lower difficulty than the High School Geometry items. As with the High School Algebra test, the High School Geometry adaptive brancher need not make any distinction between high school level and prerequisite level test items. Each Star Math High School Geometry test is administered in blocks to ensure adequate content balance.

Overarching Design Considerations

One of the fundamental Star Math High School design decisions involved the choice of how to administer the tests. The primary advantage of using computer software to administer Star Math High School tests is the ability to tailor each student's test based on his or her responses to previous items. Conventional assessments, including paper-and-pencil tests, typically entail fixed test forms: every student must respond to the same items in the same sequence. Using computer-adaptive procedures, students tend to be tested using items that appropriately match their current level of proficiency. The item selection procedures, termed Adaptive Branching, effectively customize the test for each student's achievement level, while the test proceeds.

Adaptive Branching offers significant advantages in terms of test reliability, testing time, and student motivation. Reliability improves over fixed-form tests because the test difficulty is adjusted to each individual's performance level. Most of the test items that students respond to are at levels of difficulty that closely match their achievement level. Testing time is less than that of paper-and-pencil tests, because there is no need to expose every student to a broad range of material, portions of which are inappropriate because they are either too easy for high achievers or too difficult for those with low current levels of performance.

Finally, student motivation improves simply because of these issues—test time is minimized and test content is neither too difficult nor too easy.

Another fundamental Star Math High School design decision involved the choice of the content and format of items for the test. The traditional multiple-choice format was chosen. This decision was made for interrelated reasons of efficiency, breadth of construct coverage, and objectivity and simplicity of scoring.

In Star Math High School tests, all management and test administration functions are controlled using a management system which is accessed by means of a computer with web access. This makes a number of features possible:

- Multiple schools can share a central database, such as a district-level database. Records of students transferring between schools within the district will be maintained in the database; the only information that needs revision following a transfer is the student's updated school and class assignments.
- The same database that contains Star Math High School data can contain data on other Star tests, including the Star Math Enterprise, Star Early Literacy, and Star Reading Enterprise. The Renaissance program is a powerful information management program that allows you to manage all your district, school, personnel, and student data in one place. Changes made to district, school, teacher, and student data for any of these products, as well as other Renaissance software, are reflected in every program sharing the central database.
- Multiple levels of access are available, from the test administrator within a school or classroom to teachers, principals, and district administrators.
- Renaissance takes reporting to a new level. Not only can you generate reports from the student level all the way up to the school level, but you can also limit reports to specific groups, subgroups, and combinations of subgroups. This supports "disaggregated" reporting; for example, a report might be specific to students eligible for free or reduced lunch, to English language learners, or to students who fit both categories. It also supports compiling reports by teacher, class, school, grade within a school, and many other criteria such as a specific date range. In addition, the

Renaissance consolidated reports allow you to gather data from more than one program (such as Star Math and Star Custom) at the teacher, class, school, and district level and display the information in one report.

- Since the Renaissance software is accessed through a web browser, teachers (and administrators) are able to access the program from home.

Test Interface

The Star Math High School tests' interface was designed to be both simple and effective. Students can use either the mouse or the keyboard to answer questions.

- If using the keyboard, students press one of the four letter keys (**A, B, C,** and **D**) and then press the **Enter** key (or the **return** key on Macintosh computers).
- If using the mouse, students click the answer of choice and then click **Next** to enter the answer.
- On a tablet, students tap their answer choice; then, they tap **Next**.

Practice Session

Star Math High School software includes a provision for a brief practice test preceding the test itself. The practice session allows students to get comfortable with the test interface and to make sure that they know how to operate it properly. As soon as a student has answered two out of three practice questions correctly, the program takes the student into the actual test. If the student does not successfully answer two of the three items, Star Math High School will present three more questions, and the student can pass the practice session by answering two of those questions correctly. If the student does not pass after the second attempt, the student will not proceed to the actual Star Math High School test. Even students with low math and reading skills should be able to answer the practice questions correctly. However, Star Math High School tests will halt the testing session and tell the student to ask the teacher for help if the student does not pass the practice session after the second attempt.

Students may experience difficulty with the practice questions for a variety of reasons. The student may not understand math even at the most basic level or may be confused by the "not given" response option presented in some of the practice questions. Alternatively, the student may need help using the keyboard or mouse. If this is the case, the teacher (or monitor) should help the student through the practice session during the student's next Star Math High School test. If a student still struggles with the practice questions with teacher assistance, he or she may not yet be ready to complete a Star Math High School test.

Once a student has successfully passed a practice session, the student will not be presented with practice items again on a test of the same type taken within the next 180 days.

Adaptive Branching/Test Length

Star Math's High School's branching control uses a proprietary approach somewhat more complex than the simple Rasch maximum information IRT model. The Star Math High School approach was designed to yield reliable test results for both the criterion-referenced and norm-referenced scores by adjusting item difficulty to the responses of the individual being tested while striving to minimize test length and student frustration.

In order to minimize student frustration, the first administration of the Star Math High School tests begins with items that have a difficulty level that is below what a typical student at a given grade can handle—usually one or two grades below grade placement. On the average, about 85 percent of students will be able to answer the first item correctly. On the second and subsequent administrations, the test begins about one grade lower than the ability last demonstrated within 180 days. Students generally have an 85 percent chance of answering the first item correctly on second and subsequent tests.

Test Length

Once the testing session is underway, the Star Math High School tests administer 34 items of varying difficulty based on the student's responses; this is sufficient information to obtain a reliable Scaled Score and to determine the student's math Level in either Algebra or Geometry.

The length of time needed to complete a Star Math High School tests vary across students.

Table 1 provides an overview of the testing time by grade for the students who took the Star Math High School Algebra tests during the 2018–2019 school year. The results of the analysis of test completion time indicate that half or more of students completed the test in less than 27 minutes, depending on grade, and even in the slowest grade (grade 8) 95% of students finished their Star Math High School Algebra test in less than 44 minutes.

Table 1: Average and Percentiles of Total Time to Complete the Star Math High School Algebra Assessment During the 2018–2019 School Year

Grade	Sample Size	Mean	Standard Deviation	Total Test Taking Time			
				5th Percentile	50th Percentile	95th Percentile	99th Percentile
8	11,670	26.40	9.56	12.28	25.52	43.28	52.63
9	11,050	22.30	8.32	10.37	21.32	36.67	46.15
10	4,401	21.94	8.33	9.88	21.17	36.18	43.30
11	3,478	21.93	8.26	9.98	20.98	36.38	43.33
12	1,364	20.33	7.66	8.82	19.66	33.40	40.83

Table 2 provides an overview of the Star Math High School Geometry testing time by grade for the students using data from the 2018–2019 school year. About half of the students at every grade completed the Star Math High School Geometry test in less than 30 minutes, and even in the slowest grade (grade 8) 95 percent of students finished in less than 49 minutes.

Table 2: Average and Percentiles of Total Time to Complete the Star Math High School Geometry Assessment During the 2018–2019 School Year

Grade	Sample Size	Mean	Standard Deviation	Total Test Taking Time			
				5th Percentile	50th Percentile	95th Percentile	99th Percentile
8	930	30.10	9.29	17.07	29.48	48.15	55.62
9	4,302	26.85	8.83	12.68	26.64	41.53	49.52
10	6,246	24.17	8.45	10.73	23.83	38.25	46.40
11	1,682	23.02	8.38	9.95	22.52	36.98	44.73
12	477	22.37	8.88	8.92	21.55	38.07	46.27

Test Repetition

Star Math High School score data can be used for multiple purposes such as screening, placement, planning instruction, benchmarking, and outcomes measurement.

The frequency with which the assessment is administered depends on the purpose for assessment and how the data will be used. Renaissance Learning recommends assessing students only as frequently as necessary to get the data needed. Schools that use Star for screening purposes typically administer it two to five times per year.

Star Math High School keeps track of the questions presented to each student from test session to test session and will not ask the same question more than once in any 120-day period.

Item Time Limits

The Star Math High School tests place no limits on total testing time. However, there are time limits for each test item. The per-item time limits are generous, and ensure that more than 90 percent of students can complete each item within the normal time limits. Each practice question has a 90-second time limit and each test question has a 3-minute time limit.

Standard Time Limits:

- Practice questions: 90 seconds (1.5 minutes) for each question
- Test questions 180 seconds (3 minutes) for each question

Star Math High School also provides the option of extended time limits for selected students who, in the judgment of the test administrator, require more than the standard amount of time to read and answer the test questions. Extended time limits are twice as long as standard time limits.

Extended Time Limits:

- Practice questions: 180 seconds (3 minutes) for each question
- Test questions: 360 seconds (6 minutes) for each question

Extended time may be a valuable accommodation for English language learners as well as for some students with disabilities. Test users who elect the extended time limit for their students should be aware that Star Math High School norms, as well as other technical data such as reliability and validity, are based on test administration using the standard time limits. When the extended time limit accommodation is elected, students have two times longer than the standard time limits to answer each question.

At all grades, regardless of the extended time limit setting, when a student has only 15 seconds remaining for a given item, a time-out warning appears, indicating that he or she should make a final selection and move on. Items that time out are counted as incorrect responses unless the student has the correct answer selected when the item times out. If the correct answer has been selected at that time, the item will be counted as a correct response.

If a student doesn't respond to an item, the item times out and briefly gives the student a message describing what has happened. Then the next item is presented. The student does not have an opportunity to take the item again. If a student doesn't respond to any item, all items are scored as incorrect.

Test Security

Star Math High School software includes a number of security features to protect the content of the test and to maintain the confidentiality of the test results.

Split Application Model

When students log into Star Math High School, they do not have access to the same functions that teachers, administrators, and other personnel can access. Students are allowed to take the test, but no other features available in Star Math High School are available to them; therefore, they have no access to confidential information. When teachers and administrators log in, they can manage student and class information, set preferences, and create informative reports about student test performance.

Individualized Tests

Using Adaptive Branching, every Star Math High School test consists of items chosen from a large number of items of similar difficulty based on the student's estimated ability. Because each test is individually assembled based on the student's past and present performance, identical sequences of items are rare. This feature, while motivated chiefly by psychometric considerations, contributes to test security by limiting the impact of item exposure.

Data Encryption

A major defense against unauthorized access to test content and student test scores is data encryption. All of the items and export files are encrypted. Without the appropriate decryption code, it is practically impossible to read the Star Math High School data or access or change it with other software.

Access Levels and Capabilities

Each user's level of access to a Renaissance program depends on the primary position assigned to that user. Each primary position is part of a user permission group. There are six of these groups: district level administrator, district dashboard owner, district staff, school level administrator, school staff, and teacher. By default, each user permission group is granted a specific set of user permissions; each user permission corresponds to one or more tasks that can be performed in the program. The user permissions for these groups can be changed, and user permissions can be granted or removed on an individual level.

Renaissance also allows you to restrict students' access to certain computers. This prevents students from taking Star Math High School tests from unauthorized computers (such as home computers). For more information, see https://help2.renaissance.com/setup/22509.

The security of the Star Math High School data is also protected by each person's user name (which must be unique) and password. User names and passwords identify users, and the program only allows them access to the data and features that they are allowed based on their position and the user permissions that they have been granted. Personnel who log in to Renaissance (teachers, administrators, or staff) must enter a user name and password before they can access the data and create reports. Without an appropriate user name and password, personnel cannot use the Star Math High School software.

Test Monitoring/Password Entry

Test monitoring is another useful Star Math High School security feature. Test monitoring is implemented using the Password Requirement preference, which specifies whether monitors must enter their passwords at the start of a test. Students are required to enter a user name and password to log in before taking a test. This ensures that students cannot take tests using other students' names.

Final Caveat

While Star Math High School software can do much to provide specific measures of test security, the most important line of defense against unauthorized access or misuse of the program is the user's responsibility. Teachers and test monitors need to be careful not to leave the program running unattended and to monitor all testing to prevent students from cheating, copying down questions and answers, or performing "print screens" during a test session. Taking these simple precautionary steps will help maintain Star Math High School's security and the quality and validity of its scores.

Test Administration Procedures

In order to ensure consistency and comparability of results to the Star Math High School norms, students taking Star Math High School tests should follow standard administration procedures. The testing environment should be as free from distractions for the student as possible.

The Test Administration Manual included with the Star Math High School product (https://help2.renaissance.com/US/PDF/SM/SM_TAM.pdf) describes the standard test orientation procedures that teachers should follow to prepare their students for the Star Math High School tests. These instructions are intended for use with students of all ages and were successfully field-tested with students ranging from grades 1–12. It is important to use these same instructions with all students before they take the Star Math High School tests.

Content and Item Development

Content Specification: Star Math

Content of the Star Math High School tests was developed in conjunction with content created for the broader Star Math test. All Geometry and Algebra items have been created to meet the same rigorous requirements as other Star items. These requirements include alignment to skills and state standards, content appropriateness, freedom from bias and test efficiency.

For information regarding the development of Star Math items, see "Item Development Guidelines: Star Math High School" on page 15.

Since the introduction of the initial version of the Star Math in 1998, the test has undergone a process of continuous research and improvement. Content development for the Star Math High School tests began in 2012 and continues to this day.

The Algebra and Geometry test blueprints are structured to provide a consistent assessment experience even as state specific Renaissance Learning Progressions may change, as well as the set of items associated with the blueprint. The Star Math High School test blueprints are largely fixed. Renaissance may alter the blueprint if there are data-driven reasons to make a major change to the structure of the test.

To develop the Star Math High School tests, multiple resources were consulted to determine the set of skills most appropriate for assessing the mathematics development of US high school students: typical mathematics curricula and current mathematics standards. The resources include, but are not limited to:

- US State standards from all 50 states, including the Common Core State Standards for Mathematics, updated annually
- National Mathematics Advisory Panel, *Foundations for Success: The final report of the National Mathematics Advisory Panel*
- National Council of Teachers of Mathematics (NCTM), *Principles and Standards for School Mathematics*
- *National Assessment of Educational Progress* (NAEP)

The development of the skills list for both the Algebra and Geometry tests included iterative reviews by mathematicians, mathematics educators, assessment experts, and psychometricians specializing in educational assessment. See “Appendix: Blueprint Skill Sets and Blueprint Skills” on page 63 for the Star Math High School skills.

For the purpose of content development, the skills lists for each high school test was organized to test the skills matching a typical curriculum for students taking these courses. The majority of skills tested for each test are associated with the primary test domain but skills from other domains are included if they are align to standards and curricula. The Algebra test includes skills identified for both Algebra I and Algebra II courses. Skills that have been identified as prerequisites are also included in each test.

Once skills were identified, the process for item creation and calibration began. As with all Star assessments, items are developed according to established specifications for grade-level appropriateness and then reviewed to ensure the items meet the specifications. Grade-level appropriateness is determined by multiple factors, including math skill, reading level, cognitive load, vocabulary grade level, sentence structure, sentence length, subject matter, and interest level. All item writers and editors have content-area expertise and relevant classroom experience. Those qualifications are utilized in determining grade-level appropriateness for subject matter as well as interest level. A strict development process is maintained to ensure quality item development.

Assessment items, once written, edited, and reviewed, are field tested and calibrated to estimate their Rasch difficulty parameters and goodness of fit to the model. After the initial calibration event, field testing and calibration are conducted in a single step. This is done by embedding new items in appropriate, random positions within the Star assessments to collect the item response data needed for psychometric evaluation and calibration analysis.

Following these analyses, each assessment item, along with both traditional and IRT analysis information (including fit plots) and information about the test level, form, and item identifier, are stored in an item statistics database. A panel of content reviewers then examines each item to determine whether the item meets all criteria for use in an operational assessment. More detailed information about the field testing and calibration of Star Math items may be found in the Item and Scale Calibration chapter of this manual.

The item bank for Star Math High School is dynamic. Renaissance continues to create new items for inclusion into the bank, and also periodically replaces items in order to refresh the bank.

Calculator and Formula Reference Sheets

For specific Star Math skills, a calculator or formula reference sheet is made available to the student alongside of the test item. Depending on the item and the skill addressed, either the calculator, a formula reference sheet specific to the skill, or both may be used. For the purpose of test validity, these tools are provided in the application rather than the student using their own to ensure that they are used only for appropriately identified skills.

Calculator or Formula Reference sheets are available for two general circumstances: 1) the calculation is overly difficult to perform without either a calculator or a reference chart or 2) the ability to perform the calculations is not the focus of the skill, and the calculations are difficult or time-consuming (e.g., word problems, solving equations, or finding the terms of a sequence).

Formula reference sheets are available for upper-grade skills in which the formula and math relations needed are not expected for student memorization. This decision is based on analysis of the ACT, SAT, ADP, and formula reference sheets used on state end-of-year tests.

An analysis of state assessments produced the following guidelines in determining when a calculator should be made available for Star Math:

Table 3: Determination of Calculator Availability in Star Math

Calculation	Upper Limits of Not Using a Calculator[a]
Division (1–2 step problems)	Divisors may be 1-digit, multiples of 25, fractions with 1-digit denominators, or related to basic math facts (1440/120). Other 2-digit divisors may be included if the division is carried out to only 2 or 3 places.
Multiplication (1–2 step problems)	3-digit by 2-digit, 1-digit by 4-digit (non-zero digits).
Multi-step problems (3+ steps)	2-digit by 2-digit multiplication, 1-digit divisors, other limits listed below.
Powers	2-digit numbers squared, 1-digit numbers raised to the 4th power, 2 or 3 raised to a higher power.
Square roots	Perfect squares related to square of the numbers 1–13 (e.g., square root of 144).
Nth roots	Cube roots resulting in one-digit numbers, nth roots resulting in 2 or 3.
Mean (average)	Up to 6 one- or two-digit numbers or 4 multi-digit numbers.

a. When calculation is not the focus of the skill.

Read-Aloud Audio Guidance

For students challenged by textual reading and the language involved in a Star Math test, read-aloud audio guidance was developed as an accommodation. Read-aloud guidance is turned off for all students by default, but teachers may choose to turn it on either for individual students or an entire class. The accommodation is not intended to be used for all students, blind or low-vision students, but instead is intended to assist teachers to work with students whose language skills are at a lower level than their math skills or who have reading challenges that might prevent them from understanding the item. Audio scripts are not intended to read the entire item aloud for students who cannot read or have extreme visual disabilities.

In order to ensure students receiving read-aloud audio guidance do not have an advantage over other students, some items receive a standard audio prompt of "Choose the best answer." Examples of items receiving this prompt would be if the stem included a single below-grade word such as "solve," or "simplify." Another example would be an item that includes a graphic of a coin and the student is asked to identify the value. Referring to the coin as "a quarter" in the audio prompt may make the item easier for a student who knows a quarter is worth $.25, but cannot identify the quarter visually. For content-specific scripts, only numbers and math expressions embedded within sentences are read. Audio is not included for labels on charts and graphs. Content-specific scripts will be provided for answer choices in items that would pose significant difficulty for struggling readers.

For technical reasons, a single audio file is used for each item requiring audio support, even when audio support contains both the stem and answer options. Students may replay the audio at any time, and may answer the item before the audio has finished playing.

Item Development Guidelines: Star Math High School

Star Math Algebra currently assesses 196 grade-specific blueprint skills grouped into 24 skillsets. Star Math Geometry currently assesses 95 grade-specific blueprint skills grouped into 13 skillsets. Item development is skill-specific. Each item in the item bank is developed for and clearly aligned to one skill. Answering an item correctly does not require math knowledge beyond the expected knowledge for the skill being assessed. The reading level and math level of the item are grade-level appropriate. The ATOS readability formula is used to identify reading level.

Star Math items are multiple-choice. Most items have four answer choices. An item may have two or three answer choices if appropriate for the skill. Items are distributed among difficulty levels. Correct answer choices are equally distributed by difficulty level.

Item development meets established demographic and contextual goals that are monitored during development to ensure the item bank is demographically and contextually balanced. Goals are established and tracked in the following areas: gender, ethnicity, occupation, age, and disability. Items adhere to strict bias and fairness criteria. Items are free of stereotyping, representing different groups of people in non-stereotypical settings. Items do not refer to inappropriate content that includes, but is not limited to content that presents stereotypes based on ethnicity, gender, culture, economic class, or religion; presents any ethnicity, gender, culture, economic class, or religion unfavorably; introduces inappropriate information, settings, or situations; references illegal activities; references sinister or depressing subjects; references religious activities or holidays based on religious activities; references witchcraft; or references unsafe activities.

The majority of items within a skill are homogeneous in presentation, format, or scenario, but have differing computations. A skill may have two or three scenarios which serve as the basis for homogeneous groupings of items within a skill. All items for a skill are unique. Text is typically presented as 18-point and the font used is Arial, but smaller text may be necessary to label charts or graphs. Every complete item is presented on screen with stimulus, stem and answer choices visible. Scroll bars are not used for test items to minimize cognitive load and confusion created by not having all relevant information available at once. Graphics are included in an item only when necessary to solve the problem.

Item stems meet the following criteria with limited exceptions. When possible, the stem is presented in purely mathematic form or may be limited to a single direction such as “simplify.” When an item requires more complex language, the question is concise, direct, and a complete sentence. The question is written so students can answer it without reading the distractors. Generally, completion (blank) stems are not used. If a completion stem is necessary, the stem contains enough information for the student to complete the stem without reading the distractors, and the completion blank is as close to the end of the stem as possible. The stem does not include verbal or other clues that hint at correct or incorrect distractors. The syntax and grammar are straightforward and appropriate for the grade level.

Negative construction is avoided. The stem does not contain more than one question or part. Concepts and information presented in the items are accurate, up-to-date, and verifiable. This includes but is not limited to dates, measurements, locations, and events.

Distractors meet the following criteria with limited exceptions. All distractors are plausible and reasonable. Distractors do not contain clues that hint at correct or incorrect distractors. Incorrect answers are created based on common student mistakes. Distractors that are not common mistakes may vary between being close to the correct answer or close to a distractor that is the result of a common mistake. Distractors are independent of each other, are approximately the same length, have grammatically parallel structure, and are grammatically consistent with the stem. *None of these, none of the above, not given, all of the above,* and *all of these* are generally avoided as distractors.

Star Math High School and the Renaissance Learning Progression for Math

Star Math High School bridges assessment and instruction through a research-based Renaissance learning progression for math to help teachers make effective instructional decisions and to adjust instruction to meet the needs of students at different achievement levels. All 50 US states have their own individualized learning progression, which is based on their own state's standards and updated yearly as standards change. The Renaissance learning progression for math identifies the continuum—or instructional sequence—of math concepts written as skills spanning from early numeracy through high-school level algebra and geometry. It was developed in consultation with leading experts in mathematics and supported by calibration data and psychometric analysis.

Item and Scale Calibration

Background

The Star Math High School tests of Algebra and Geometry mark a fourth phase in the evolution of Star Math. Star Math versions 1 and 2 were tests of general math achievement used in grades 1–12. The Enterprise version, introduced in 2011, is a standards-based math achievement test for the same grade range. All of those Star Math versions have reported scores on the same vertical scale, a single common scale which spans the entire grade range.

The Star Math High School tests mark a change from the general math assessments, in two significant respects.

1. Their content is limited to Algebra or Geometry skills typically taught in the high school or middle school grades.
2. Its test items are grade agnostic, and the tests are intended for use at whatever grade(s) the subject matter is taught. For example, some students may take Algebra 1 in grade 9, while others may take it in grade 7 or 8, or even earlier or later.

To facilitate the comparability of scores on the general math, algebra and geometry versions of Star Math High School, the original Star Math scale was adopted for use in calibrating the difficulty of the test items to be used in the Algebra and Geometry tests and in reporting Algebra and Geometry test scores. This chapter describes the process used to accomplish that.

That process is an application of item response theory—specifically the Rasch 1-parameter logistic response model that has been applied to all earlier Star Math versions. Star Math tests use the Rasch model to calibrate item difficulty, as well as to calculate students' test scores. The sections below describe the Rasch model in general terms, then provide details about the calibration of the Star Math High School Algebra and Geometry items on the established Star Math vertical score scale.

Item calibration entails estimating the scaled difficulty of test items by administering them to examinees whose ability is known or estimated, then fitting item response models that express the probability of a correct response to each item as a function of examinee ability. To provide accurate item difficulty parameter estimates requires an adequate number of responses to each item, from examinees spanning a broad range of ability. The distribution of ability in the examinee samples need not be closely representative of the distribution of ability in the population, but it needs to be diverse, with large enough numbers of observations above and below the middle of the ability range, as well as from the middle itself.

The Rasch Item Response Model

The Star Math calibration data are analyzed using item response theory (IRT) methods.

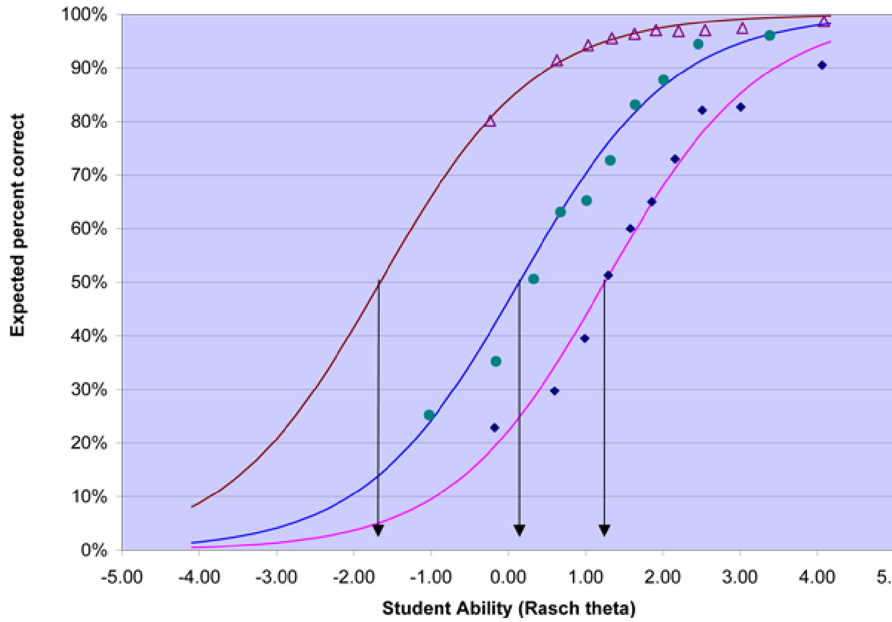
With IRT, the performance of students and the difficulty of test items are expressed on the same scale. To accomplish this, every test question is calibrated. Calibration is an IRT-based analytical method for estimating the location of a test question on the scale used to measure examinee ability. It is done by administering each question to hundreds and sometimes thousands of students with known performance levels. As a result of calibration, Star “knows” the relative difficulty of every item from kindergarten through grade 12, and expresses it on a developmental scale spanning from the easiest to the hardest questions in the item bank. After taking a Star assessment, a student’s score can be plotted on this developmental scale. Placing students and items on the same scale makes it possible to assign scores on the same scale even though students take different tests. IRT also provides a means to estimate what skills a student knows and doesn’t know, without explicitly testing each and every skill.

IRT methods develop mathematical models of the relationship of student ability to the difficulty of specific test questions; more specifically, they model the probability of a correct response to each test question as a function of student ability. Although IRT methods encompass a family of mathematical models, the one-parameter (or Rasch) IRT model was selected for the Star Math High School data both for its simplicity and for its ability to accurately model the performance of the Star Math High School items.

Within IRT, the probability of answering an item correctly is a function of the student’s ability and the difficulty of the item. Since IRT places the item difficulty and student ability on the same scale, this relationship can be represented graphically in the form of an item response function (IRF).

Figure 1 on page 20 is a plot of three item response functions: one for an easy item, one for a more difficult one, and one for a very difficult item. Each plot is a continuous S-shaped (ogive) curve. The horizontal axis is the scale of student ability, ranging from very low ability (–5.0 on the scale) to very high ability (+5.0 on the scale). The vertical axis is the percent of students expected to answer each of the three items correctly at any given point on the ability scale. Notice that the expected percent correct increases as student ability increases, but varies from one item to another.

Figure 1: Three Examples of Item Response Functions



In Figure 1, each item's difficulty is the scale point where the expected percent correct is exactly 50. These points are depicted by vertical lines going from the 50% point to the corresponding locations on the ability scale. The easiest item has a difficulty scale value of about –1.67; this means that students located at –1.67 on the ability scale have a 50-50 chance of answering that item right. The scale values of the other two items are approximately +0.20 and +1.25, respectively.

Calibration of test items estimates the IRT (Rasch) difficulty parameter for each test item and places all of the item parameters onto a single scale used to assess the difficulty of test items, and the ability of students, ranging from Kindergarten through 12th grade level. The difficulty parameter for each item is estimated, along with measures to indicate how well the item conforms to (or "fits") the theoretical expectations of the presumed IRT model.

Also plotted in Figure 1 are the actual percentages of correct responses of groups of students to all three items. Each group is represented as a small triangle, circle, or diamond. Each of those geometric symbols is a plot of the percent correct against the average ability level of the group. Ten groups' data are plotted for each item; the triangular points represent the groups responding to the easiest item. The circles and diamonds, respectively, represent the groups responding to the moderate and to the most difficult item.

Calibration of Items for Star Math High School Algebra and Geometry

Multiple fixed-form field tests served as the vehicle for new item response data collection; the analyses themselves fit response models to the new items, using the response data itself as the basis for estimating examinee ability.

Calibration of items for the Star Math High School Algebra and Geometry tests took place in three successive phases. The first phase took place in the 2011–12 school year, with a total of 1,334 items administered to an overall sample of 20,858 cases. In the second phase, 925 items were administered, in school year 2012–2013. The third phase also took place in 2013, with 603 items administered. Within each phase, students took field test forms consisting of mostly Algebra 1, Algebra 2, or Geometry items, depending on the courses of study the students were enrolled in. Regardless of the subject, each field test form included several Star Math general content items that had previously been calibrated; such items were used as “scale anchors,” to facilitate ultimately expressing each new Algebra or Geometry Rasch difficulty parameter on the established Star Math Rasch scale.

Phase 1

The Phase 1 field test included more than 20 thousand students in Grades 9 through 11, responding to approximately 120 new test forms incorporating more than 1,300 test items including new items as well as some pre-calibrated general Star Math items intended to anchor the IRT Rasch difficulty scale. Grade 9 students took forms containing Algebra 1 and Star Math anchor items; Grade 10 students took forms containing Geometry and Star Math anchor items; Grade 11 students took forms containing Algebra 2 and Star Math anchor items. More than 300 students responded to each of the test forms.

The intent of the project was to calibrate the new test high school items on the existing Star Math Rasch difficulty scale, using project-specific anchor items to anchor test forms within each content area, and the existing Star Math items to anchor the items from each content area to the existing Star Math Rasch scale.

Field Test Design

Phase 1 field test forms were developed for the three subjects: Algebra 1, Algebra 2 and Geometry. There were 14 forms for each subject. Each form had three variants in which the order of items was altered. The total number of variant forms per course was 42 (14 times 3). Each variant form had 43 items; 3 practice items, 30 items for calibration and 10 anchor items.

Across the forms in each subject, there were 40 anchor items. Half of those anchor items were a common set of Star Math 2 anchor items. The other half of the anchor items were operational or mastery items in related content. A total of 1,334 items were administered. Of those, 9 were practice items, 66 were common anchor items, 420 were Algebra 1 calibration items, 420 were Algebra 2 calibration items and 419 were Geometry calibration items. 20,858 students were tested: 7,388 in Algebra 1, 6754 in Algebra 2 and 6716 in Geometry.

Although a nationally representative examinee sample is not required for item calibration, it is useful to evaluate the diversity of the samples who contributed to the calibration data. Table 4, Table 5, and Table 6 summarize some characteristics of the Phase 1 students.

Table 4 displays numbers of students taking field tests of the Algebra 1, Geometry, or Algebra 2 items in each of grades 9, 10 and 11. It shows that the number of students field testing in each subject was nearly equal—approximately 6,700 to 7,400 students per subject, with similar numbers of students from each grade. By far the majority of students taking Algebra 1 were 9th graders; 10th graders accounted for the majority of Geometry students; and 11th graders predominantly took Algebra 2 tests.

In all, 20,858 students, in US schools participated in the calibration field test study; selected demographic data on the U.S. students are in the following tables.

Table 4: Sample by High School Grade—Phase 1 Spring 2011 (N = 20,858 US Students)

Course	Grade 9	Grade 10	Grade 11
Algebra 1	7,363	15	10
Geometry	27	6,665	24
Algebra 2	7	16	6,731
Total	7,397	6,696	6,765

Table 5 displays the recorded ethnicity of those examinees, and provides population percentages that indicate that the race/ethnic distribution of the field test examinees was similar to the distribution in the US national population at about the same time.

Table 5: Sample Race/Ethnicity, Star Math High School Calibration Study—Phase 1, Spring 2011 (N = 20,858 US Students)

Race/Ethnicity Description	Observed Percentage	Population Percentage
American Indian or Alaskan Native	5.01	1.1
Asian or Pacific Islander	5.34	3.9
Black	12.78	16.8
Hispanic	12.94	14.7
Other Race or Ethnicity	0.00	—
White	68.54	63.5

Table 6 displays the distribution of the examinees by gender. As the tables shows, the distribution included nearly equal numbers of female and male students.

Table 6: Sample by Gender, Star Math High School Calibration Study—Phase 1 Spring 2011 (N = 20,858 US Students)

Gender	Observations	Observed Percentage	Population Percentage
Female	10,391	50.51	48.39
Male	10,181	49.49	51.61
Total	20,572		
Missing	286		

Phase 2

A second phase of calibration of high school test items took place in 2012; in that phase, 9,284 tests were administered. The Phase 2 field test again included students in Grades 9 through 11, responding to 144 new test forms incorporating a total of 925 items, including 30 of the pre-calibrated Star Math scale anchor items employed in Phase 1. As before, Grade 9 students took forms containing Algebra 1 and Star Math anchor items; Grade 10 students took forms containing Geometry and Star Math anchor items; Grade 11 students took forms containing Algebra 2 and Star Math anchor items.

Field Test Design

The field test included 888 items to be calibrated, 30 scale anchor items, and 7 practice items. All these items were distributed among 144 forms: 3 variants of 48 unique forms, with items arranged in a different order within each variant. Phase 2 test forms consisted of 33 items each. The first 3 questions of each form were practice items. The next 30 were a combination of 25 to-be-calibrated items and 5 anchor items per test.

A total of 925 items, including 7 practice items, were calibrated in Phase 2. In addition to the 30 anchor items, there were 660 items administered in the Algebra tests, and 327 items in the Geometry tests; 86 items were common to both the Algebra and Geometry tests.

Phase 3

The Phase 3 field test included a total of 12,412 student tests administered in Grades 9 through 11, responding to 90 new test forms. In all, the field test forms incorporated more than 390 new test items, plus approximately 180 items carried over from Phase 2, and the same 30 pre-calibrated Star Math scale anchor items employed in Phase 2. Grade 9 students took forms containing Algebra 1 and Star Math anchor items; Grade 10 students took forms containing Geometry and Star Math anchor items; Grade 11 students took forms containing Algebra 2 and Star Math anchor items. More than 300 students responded to each of the test forms.

Field Test Design

The field test included 571 items to be calibrated, 30 scale anchor items, and 7 practice items; there were 442 Algebra items and 129 Geometry items.

All these items were distributed among 90 forms: 3 variants of 30 unique forms, with items arranged in a different order within each variant. Phase 3 test forms consisted of 33 items each. The first 3 questions of each form were practice items. The next 30 were a combination of 25 to-be-calibrated items and 5 anchor items per test.

Item Analyses

Both traditional and IRT item analyses were conducted for the item response data collected. The traditional analyses yielded proportion correct statistics, as well as biserial and point-biserial correlations between scores on the new items and actual scores on the Star Math tests.

All Star Math calibration analyses since 2008 followed similar courses. Following extensive quality control checks, the item response data are analyzed using both traditional item analysis techniques and item response theory (IRT) methods. For each test item, the following information is derived using traditional psychometric item analysis techniques:

- The number of students who attempted to answer the item.
- The number of students who did not attempt to answer the item.
- The percentage of students who answered the item correctly (a traditional measure of difficulty).
- The percentage of students answering each alternative option.
- The correlation between answering the item correctly and the total score (a traditional measure of discrimination).
- The correlation between the endorsement of each alternative answer and the total score.

Traditional Item Difficulty

The difficulty of an item in traditional item analysis is the percentage (or proportion) of students who answer the item correctly. This is typically referred to as the "p-value" of the item. Low p-values (such as 15%) indicate that the item is difficult since only a small percentage of students answered it correctly. High p-values indicate that the majority of students answered the item correctly and thus, the item is easy. It should be noted that the p-value only has meaning for a particular item relative to the characteristics of the sample of students who responded to it.

Item Discriminating Power

The traditional measure of the discriminating power of an item is the correlation between the "score" on the item (correct or incorrect) and the total test score. Items that correlate highly with total test score will also tend to correlate with one another more highly and produce a test with more internal consistency. For the correct answer, the higher the correlation between the item score and the total score, the better the item is at discriminating between low-scoring and high-scoring individuals. When the correlation between the correct answer and the total test is low (or negative), the item is most likely not performing as intended. The correlation between endorsing incorrect answers and the total score should generally be negative, since there should not be a positive relationship between selecting an incorrect answer and scoring higher on the overall test.

At least two different correlation coefficients are commonly used during item analysis: the point-biserial and the biserial coefficients. The former is a traditional product-moment correlation that is readily calculated, but is known to be somewhat biased in the case of items with p-values that deviate from 0.50. The biserial correlation is derived from the point-biserial and the p-value, and is preferred by many because it in effect corrects for the point-biserial's bias at low and high p-values. For item analysis of Star Math High School data, the correlation coefficient of choice was the biserial.

Urry (1975) demonstrated that in cases where items could be answered correctly by guessing (e.g., multiple choice items) the value of the biserial correlation is itself attenuated at p-values approaching the chance level. He derived a correction for this attenuation, which we will refer to as the "Urry biserial correlation." Urry demonstrated that multiple choice adaptive tests are more efficient than conventional tests only if the adaptive tests use items with Urry biserial values that are considerably higher than the target levels often used to select items for conventional test use. His suggestion was to reject items with Urry biserial values lower than 0.62. Item analyses of the Star Math High School have used the Urry biserial as the correlation coefficient of choice; item selection/rejection decisions have been based in part on his suggested criterion of 0.62 or higher.

Rules for Item Retention

Following these analyses, each test item, along with both traditional and IRT analysis information (including the item response function (IRF) and empirical item response function (EIRF) plots), and information about the test level, form, and item identifier is stored in a specialized item statistics database system. A panel of internal reviewers then examines each item's statistics to determine whether the item met all criteria for inclusion in the bank of Star Math High School items. The item statistics database system allows experts easy access to all available information about an item in order to interactively designate items that, in their opinion, meet acceptable standards for inclusion in the Star Math item bank.

Generally, items are eliminated when they meet one or more of the following criteria:

- One or more incorrect answer option has a positive item discrimination value;

- Sample size of students responding to the item less than 1000;
- The traditional item difficulty indicated that the item was too difficult or too easy;
- Item-total correlation (item discrimination) less than the minimum (Urry biserial => 0.62);
- The item does not appear to fit the Rasch IRT model.

In the case of the batch of 2,473 items used in the example of Star Math item calibration above, 884 items (36%) met all the retention rules above, and were accepted for operational use as part of the Star Math adaptive test item bank. Another 538 items met all criteria except the Urry biserial target. Such items would meet commonly applied criteria for use in most conventional tests; those 538 items were retained for use for certain analytical purposes, but will not be used for adaptive testing in Star Math.

The Star Math High School Item Bank

After applying the item selection rules, 598 of the items calibrated in Phases 1 through 3 were selected for use in the High School Algebra test, and 362 were selected for use in High School Geometry. To supplement the newly calibrated items, 370 previously calibrated Star Math items were chosen for use in the Algebra test, and 166 such items were chosen for us in the Geometry test. In all, the Star Math High School Algebra item bank contains 968 items, and the Star Math High School Geometry item bank contains 528. Some items were selected for use in both tests.

Computer-Adaptive Test Design

In computer-adaptive tests, such as the Star Math High School Algebra and Geometry tests, the items taken by a student are dynamically selected in light of that student’s performance during the testing session. Thus, the assessment for a low-performing student’s knowledge of math operations may branch to easier operations to better estimate math achievement level, and the assessment for high-performing students may branch to more challenging operations or concepts to better determine the breadth of their math knowledge and their math achievement level.

During an adaptive test, a student may be “routed” to items at the lowest level of difficulty within the overall pool of items, dependent upon the student’s unfolding performance during the testing session. In general, when an item is answered correctly, the student is routed to a more difficult item. When an item is answered incorrectly, the student is instead routed to an easier item. In the case of the two Star Math High School tests, the brancher targets for selection items with a 67 percent expectation of a correct response, given the student's estimated ability, and the item's calibrated difficulty.

Both Star Math Algebra and Geometry tests are fixed-length, 34-item adaptive tests. Students who have not taken a Star Math test within 180 days initially receive an item whose difficulty level is relatively easy for students at that grade level. This minimizes any effects of initial anxiety that students may have when Starting the test and serves to better facilitate the students' initial reactions to the test. The starting points vary by grade level and are based on research conducted as part of the norming process.

When a student has taken a Star Math test within the previous 180 days, the appropriate starting point is based on his or her previous test score information. Following the administration of the initial item, and after the student has entered an answer, the program determines an updated estimate of the student's math achievement level. Then, it selects the next item randomly from among all of the available items having a difficulty level that closely match a target based on the estimated achievement level. Randomization of items with difficulty values near the target level allows the program to avoid overexposure of test items. Items that have been administered to the same student within the past 120 days are not available for administration. Because the item pools make a large number of items available for selection, these minor constraints have a negligible impact on the quality of each Star Math High School computer-adaptive test.

Items that have been administered to the same student within the past 120 days are not available for administration. In addition, to avoid frustration, items that are intended to measure advanced mathematical concepts and operations that are more than three grade levels beyond the student's grade level, as determined by where such concepts or operations are typically introduced in math textbooks, are also not available for administration. Because the item pools make a large number of items available for selection, these minor constraints have a negligible impact on the quality of each Star Math computer-adaptive test.

Scoring in the Star Math High School Tests

Following the administration of each Star Math High School item, and after the student has selected a response, an updated estimate of the student's underlying math achievement level is computed based on the student's responses to all of the items administered up to that point. A proprietary Bayesian-modal item response theory estimation method is used for scoring until the student has answered at least one item correctly and at least one item incorrectly. Once the student has met this 1-correct/1-incorrect criterion, the software uses a proprietary Maximum-Likelihood IRT estimation procedure to avoid any potential bias in the Scaled Scores.

This approach to scoring enables the software to provide scaled scores that are statistically consistent and efficient. Scaled scores are expressed on a common scale that spans all grade levels covered by the Star Math High School tests. Because the software expresses scaled scores on a common scale, scaled scores are directly comparable with each other, regardless of grade level. Norm-referenced scores, such as percentile ranks and grade equivalent (GE) scores are not reported for the Star Math High School Algebra or Geometry test.

A New Scale for Reporting Star Math High School Test Scores

In 1998, Renaissance released the initial 24-item version of Star Math. In 2011, the 34-item standards-based Star Math Enterprise test was published. Star Math High School tests are on the same scales as Star Math Enterprise that spans grades K–12. Although Star Math measures constructs that are different from those assessed in Star Reading, a common scale - the Unified Score Scale - that can be used to report scores on both tests was developed in 2015. The Unified Score Scale was introduced into use in the 2017–2018 school year as an optional alternative scale for reporting achievement on all Star tests.

The Unified Score Scale is derived from the Star Reading Rasch scale of ability and difficulty, which was first introduced with the development of Star Reading Version 2. The Unified Star Math scale was developed by performing the following steps:

- The Rasch scale used by Star Math Enterprise was linked (transformed) to the Star Reading Enterprise Rasch scale.
- A linear transformation of the transformed Rasch scale was developed that spans the entire range of knowledge and skills measured by Star Math.

Details of these two steps are presented below.

1. The Rasch scale used by Star Math Enterprise was linked to the Star Reading Enterprise Rasch scale.

   In this step, a linear transformation of the Star Math Rasch scale to the Rasch scale used by Star Reading was developed, using a method for linear equating of IRT (item response theory) scales described by Kolen and Brennan (2004, pages 162–165).

2. Because Rasch scores are expressed as decimal fractions, and may be either negative or positive, a more user-friendly scale score was developed that uses positive integer numbers only. A linear transformation of the extended Star Reading Rasch scale was developed that spans the entire range of knowledge and skills measured by both Star Math and Star Reading. The transformation formula is as follows:

   Unified Scale Score = INT (42.93 * Star Reading Rasch Score + 958.74)

   where the Star Reading Rasch score has been extended downwards to values as low as –20.00.

   Following are some features and considerations in the development of that scale, called here the "Unified scale."

   a. For both Star Math and Star Reading, the range of reported unified scales range is from 600–1400. Anchor points were chosen such that the unified scale score of 600 is approximately equivalent to a Star Math scale score of 0, and a unified score of 1400 is the approximate equivalent of 1300 on the Star Math scale.

   b. The scale is extensible upwards and downwards. Currently, the highest reported Star Math unified scale is 1400; but there is no theoretical limit: If Star Math content were extended beyond the high school level, the range of the new scale can be extended upward without limit, as needed. The lowest point is now set at 600; but the unified scale can readily be extended downward as low as 0, if a reason arises to do so.

Table 7 contains a table of selected Star Math Rasch ability scores and their equivalents on the Star Math and Unified Score scales.

Table 7: Some Star Math Rasch Scores and their Equivalents on the Star Math and Unified Score Scales

Minimum Rasch Score	Star Math Scaled Score	Unified Scale Score
–8.35	0	600
–7.72	50	638
–7.08	100	668
–6.45	150	699
–5.81	200	730
–5.18	250	761
–4.54	300	791
–3.91	350	822
–3.27	400	853
–2.64	450	884
–2.00	500	914
–1.37	550	945
–0.74	600	976
–0.10	650	1007
0.54	700	1037
1.17	750	1068
1.81	800	1099
2.44	850	1130
3.07	900	1160
3.71	950	1191
4.34	1000	1222
4.98	1050	1253
5.61	1100	1283
6.25	1150	1314
6.88	1200	1345
7.52	1250	1376
8.15	1300	1400

Reliability and Measurement Precision

Measurement is subject to error. A measurement that is subject to a great deal of error is said to be imprecise; a measurement that is subject to relatively little error is said to be *reliable*. In psychometrics, the term *reliability* refers to the degree of measurement precision, expressed as a proportion. A test with perfect score precision would have a reliability coefficient equal to 1, meaning that 100 percent of the variation among persons' scores is attributable to variation in the attribute the test measures, and none of the variation is attributable to error. Perfect reliability is probably unattainable in educational measurement; for example, a test with a reliability coefficient of 0.90 is much more likely. On such a test, 90 percent of the variation among students' scores is attributable to the attribute being measured, and 10 percent is attributable to errors of measurement. Another way to think of score reliability is as a measure of the consistency of test scores. Two kinds of consistency are of concern when evaluating a test's measurement precision: internal consistency and consistency between different measurements.

First, internal consistency refers to the degree of confidence one can have in the precision of scores from a single measurement. If the test's internal consistency is 95 percent, just 5 percent of the variation of test scores is attributable to measurement error.

Second, reliability as a measure of consistency between two different measurements indicates the extent to which a test yields consistent results from one administration to another and from one test form to another. Tests must yield somewhat consistent results in order to be useful; this reliability coefficient is obtained by calculating the correlation coefficient between students' scores on two different occasions, or on two alternate versions of the test given at the same occasion.

Because the amount of the attribute being measured may change over time, and the content of tests may differ from one version to another, the internal consistency reliability coefficient is generally higher than the correlation between scores obtained on different administrations.

There are a variety of methods of estimating the reliability coefficient of a test. Methods such as Cronbach's alpha and split-half reliability are single administration methods and assess internal consistency. Coefficients of correlation calculated between scores on alternate forms, or on similar tests administered two or more times on different occasions, are used to assess alternate forms reliability, or test-retest reliability (stability).

In a computerized adaptive test such as Star Math High School Algebra and Geometry tests, content varies from one administration to another, and it also varies with each student's performance. Another feature of computerized adaptive tests based on Item Response Theory (IRT) is that the degree of measurement error can be expressed for each student's test individually.

The Star Math High School Algebra and Geometry tests provide two ways to evaluate the reliability of scores: reliability coefficients, which indicate the overall precision of a set of test scores, and conditional standard errors of measurement (CSEM), which provide an index of the degree of error in an individual test score. A reliability coefficient is a summary statistic that reflects the average amount of measurement precision in a specific examinee group or in a population as a whole. In Star Math High School Algebra and Geometry, the CSEM is an estimate of the unreliability of each individual test scores. While a reliability coefficient is a single value that applies to the test in general, the magnitude of the CSEM may vary substantially from one person's test score to another's.

This chapter presents three different types of reliability coefficients: generic reliability, split-half reliability, and alternate forms (test-retest) reliability. This is followed by statistics on the conditional standard error of measurement of Star Math High School Algebra and Geometry test scores.

The reliability and measurement error presentation is divided into two sections below:

1. Reliability
2. Measurement error

The reliability coefficients and conditional standard errors of measurement are presented for scores expressed on both the Enterprise Star Math High School Algebra and Geometry Scale and the recently developed Star Unified Scale.

Star Math High School Algebra and Geometry Tests

Generic Reliability

Test reliability is generally defined as the proportion of test score variance that is attributable to true variation in the trait the test measures. This can be expressed analytically as:

$$\textit{Reliability} = 1 - \frac{\sigma^2_{error}}{\sigma^2_{total}}$$

where σ^2_{error} is the variance of the errors of measurement, and σ^2_{total} is the variance of test scores. In Star Math High School Algebra and Geometry tests, the variance of the test scores is easily calculated from Scaled Score data. The variance of the errors of measurement may be estimated from the conditional standard error of measurement (CSEM) statistics that accompany each of the IRT-based test scores, including the Scaled Scores, as depicted below.

$$\sigma^2_{error} = \frac{1}{n}\sum_{i=1}^{n} SEM_i^2$$

where the summation is over the squared values of the reported CSEM for students i = 1 to n. In each Star Math High School Algebra and Geometry test, CSEM is calculated along with the IRT ability estimate and Scaled Score. Squaring and summing the CSEM values yields an estimate of total squared error; dividing by the number of observations yields an estimate of mean squared error, which in this case is tantamount to error variance. “Generic” reliability is then estimated by calculating the ratio of error variance to Scaled Score variance, and subtracting that ratio from 1.

Data for the reliability estimates came from the Star Math High School Algebra and Geometry test score data from 2018–2019 academic year (August 2018–July 2019). Although the Algebra and Geometry tests were given to advanced 7th and 8th graders in middle school as well as to high school students, these two lowers grades were excluded from our analysis.

To ensure that data was representative of the students taking the tests, sampling was conducted at the overall and grade specific levels for the calculations of the reliability estimates.

The results of the reliability analyses are presented in Table 8 and Table 9, in both the Unified and Enterprise scale. Since the method for estimating the generic reliability is not susceptible to error variance introduced by repeated testing, multiple occasions, and alternate forms, the resulting estimates of reliability are generally higher than the more conservative alternate forms reliability coefficients. These generic reliability coefficients are, therefore, plausible upper-bound estimates of the internal consistency reliability of the Star Math High School Algebra and Geometry computer-adaptive test.

Table 8: Reliability Estimates from the Star Math High School Algebra 2018–2019 Data on both the Unified Scale and the Enterprise Scale

Grade	Algebra Reliability Estimates									
	Generic			Split-Half			Test-Retest			
	N	ρxx (Unified)	ρxx (Enterprise)	N	ρxx (Unified)	ρxx (Enterprise)	N	ρxx (Unified)	ρxx (Enterprise)	Days Between
9	3,000	0.89	0.89	3,000	0.89	0.89	500	0.80	0.80	114
10	4,500	0.91	0.91	4,500	0.91	0.91	1,000	0.83	0.83	101
11	4,000	0.90	0.90	4,000	0.90	0.90	500	0.81	0.81	108
12	1,500	0.91	0.91	1,500	0.91	0.91	350	0.79	0.79	104
Overall	13,000	0.91	0.91	13,000	0.91	0.91	2,350	0.83	0.83	106

Table 9: Reliability Estimates from the Star Math High School Geometry 2018–2019 Data on both the Unified Scale and the Enterprise Scale

Grade	Geometry Reliability Estimates									
	Generic			Split-Half			Test-Retest			
	N	ρxx (Unified)	ρxx (Enterprise)	N	ρxx (Unified)	ρxx (Enterprise)	N	ρxx (Unified)	ρxx (Enterprise)	Days Between
9	4,000	0.87	0.87	4,000	0.87	0.87	1,000	0.74	0.74	127
10	5,000	0.86	0.86	5,000	0.86	0.86	2,000	0.71	0.71	120
11	1,500	0.85	0.85	1,500	0.86	0.86	500	0.73	0.73	132
12	500	0.83	0.83	500	0.84	0.84	100	0.73	0.73	105
Overall	11,000	0.87	0.87	11,000	0.87	0.87	3,600	0.74	0.74	123

Table 8 presents the generic reliability estimates for Algebra on both the Unified score scale and the Enterprise score scale. Because both the Unified scaled and the Enterprise scale are linear transformations of the underlying Rasch scores, the reliability estimates are the same across both scales. Results in Table 8 indicate that for Algebra, the overall generic reliability of the scores is about 0.91. Coefficients range from a low of 0.89 in grade 9 to a high of 0.91 in grades 10 and 12.

Table 9 presents the generic reliability estimates for Geometry, the overall generic reliability of the scores is about 0.87. Coefficients range from a low of 0.83 in grade 12 to a high of 0.87 in grade 9.

As the data in Table 8 and Table 9 show, generic reliability of both Star Math High School Algebra and Geometry is quite high, grade by grade and overall. Star Math High School Algebra and Geometry also demonstrate fairly high test-retest consistency as shown in the rightmost columns of the same table. The technical quality of both Star Math High School Algebra and Geometry is quite high.

Split-Half Reliability

While generic reliability does provide a plausible estimate of measurement precision, it is a theoretical estimate, as opposed to traditional reliability coefficients, which are more firmly based on item response data. Traditional internal consistency reliability coefficients such as Cronbach's alpha and Kuder-Richardson Formula 20 (KR-20) are not applicable to adaptive tests. However, an estimate of internal consistency reliability can be calculated using the split-half method.

A split-half reliability coefficient is calculated in three steps. First, the test is divided into two halves, and scores are calculated for each half. Second, the correlation between the two resulting sets of scores is calculated; this correlation is an estimate of the reliability of a half-length test. Third, the resulting reliability value is adjusted, using the Spearman-Brown formula, to estimate the reliability of the full-length test.

In internal simulation studies, the split-half method provided accurate estimates of the internal consistency reliability of adaptive tests, and so it has been used to provide estimates of Star Math High School Algebra and Geometry reliability. These split-half reliability coefficients are independent of the generic reliability approach discussed earlier and more firmly grounded in the item response data. Split- half scores were based on all of the 34 items of the Star Math High School Algebra and Geometry tests; scores based on the odd- and the even-numbered items were calculated separately. The correlations between the two sets of scores were corrected to a length of 34 items, yielding the split-half reliability estimates displayed in Table 8 and Table 9.

Results indicate that the overall split-half reliability of scores is 0.91 for Algebra and 0.87 for Geometry. The coefficients range from a low of 0.89 in grade 9 to a high of 0.91 in grades 10 and 12 for Algebra, and a low of 0.84 in grade 12 to a high of 0.87 in grade 9 for Geometry. These reliability estimates are quite consistent across grades 9–12, and quite high, again a result of the measurement efficiency inherent in the adaptive nature of the Star Math High School Algebra and Geometry test.

Alternate Form Reliability

Another method of evaluating the reliability of a test is to administer the test twice to the same examinees. A reliability coefficient is obtained by calculating the correlation between the two sets of test scores. This is called a test-retest reliability coefficient if the same test was administered both times and an alternate forms reliability coefficient if different, but parallel, tests were used.

Content sampling, temporal changes in individuals' performance, and growth or decline over time can affect alternate forms reliability coefficients, usually making them appreciably lower than internal consistency reliability coefficients.

The alternate forms reliability study provided estimates of Star Math High School Algebra and Geometry reliability using a variation of the test-retest method. In the traditional approach to test- retest reliability, students take the same test twice, with a short time interval, usually a few days, between administrations. In contrast, the Star Math High School Algebra and Geometry alternate form reliability study administered two different tests, and avoiding during the second test the use of any items the student had encountered in the first test. All other aspects of the two tests were identical. The correlation coefficient between the scores on the two tests was taken as the reliability estimate.

The alternate form reliability estimates for the Star Math High School Algebra and Geometry test were calculated using both the Star Math High School Algebra and Geometry Unified scaled scores and the Enterprise scaled scores. Checks were made for valid test data on both test administrations and cases of apparent motivational discrepancies were removed.

Table 8 and Table 9 include overall and within-grade alternate reliability, along with an indication of the average number of days between testing occasions. The average number of days between testing occasions ranged from 101–114 days for Algebra and 105–132 days for Geometry. Results indicate that the overall alternate test reliability of the scores is about 0.83 for Algebra and 0.74 for Geometry. The alternate form coefficients range from a low of 0.79 in grade 12 to a high of 0.83 in grade 10 for Algebra; for Geometry, the coefficients range from a low of 0.71 of grade 10 to a high of 0.74 in grade 9.

Because errors of measurement due to content sampling and temporal changes in individuals' performance can affect this correlation coefficient, this type of reliability estimate provides a conservative estimate of the reliability of a single Star Math High School Algebra or Geometry administration. In other words, the actual Star Math High School Algebra and Geometry reliability is likely higher than the alternate form reliability estimates indicate.

A general rule of thumb is that a reliability estimate equal to or greater than 0.80 is acceptable for making instructional decisions about students; the generic reliability estimates in Table 8 indicate an acceptable range of score reliability for both the Algebra and the Geometry tests. The test-retest reliability estimates are lower than the generic reliability estimates and somewhat constant across all grades. The test-retest reliability estimate is based on relatively smaller sample sizes.

Standard Error of Measurement

When interpreting the results of any test instrument, it is important to remember that the scores represent estimates of a student's true ability level. Test scores are not absolute or exact measures of achievement performance since the assessment is based on a sample of items drawn from a content domain. Nor is a single test score infallible in the information that it provides. The standard error of measurement can be thought of as a measure of the precision of a given score. The standard error of measurement describes the extent to which scores would be expected to fluctuate because of chance factors. If measurement errors follow a normal distribution, an SEM of 18 means that if a student were tested repeatedly, his or her scores would fluctuate within plus or minus18 points of his or her first score about 68 percent of the time, and within 36 points (twice the SEM) roughly 95 percent of the time. Since reliability can also be regarded as a measure of precision, there is an inverse relationship between the reliability of a test and the standard error of measurement for the scores it produces: as reliability increases, standard error of measurement decreases.

The Star Math High School Algebra and Geometry tests differ from traditional tests in at least two respects with regard to the standard error of measurement. First, Star Math High School Algebra and Geometry software computes the SEM for each individual student based on his or her performance, unlike most traditional fixed tests that report the same SEM value for every examinee for a linear fixed form. Each administration of Star Math High School Algebra and Geometry yields a unique "conditional" SEM (CSEM) that reflects the amount of information estimated to be in the specific combination of items that a student received in his or her individual test. Second, because the Star Math High School Algebra and Geometry tests are adaptive, the CSEM will tend to be lower than that of a conventional test, particularly at the highest and lowest score levels, where conventional tests' measurement precision is weakest. Because the adaptive testing process attempts to provide equally precise measurement, regardless of the student's ability level, the average CSEMs for the IRT ability estimates are very similar for all students.

Table 10 and Table 11 contain two different sets of estimates of Star Math High School Algebra and Geometry measurement error, respectively: conditional standard error of measurement (CSEM) and global standard error of measurement (SEM). Conditional SEM was just described; the estimates of CSEM in Table 10 and Table 11 are the average CSEM values observed for each grade.

Global standard error of measurement is based on the traditional SEM estimation method, using internal consistency reliability and the variance of the test scores to estimate the SEM:

$SEM = SQRT(1 - \rho)\ \sigma_x$

where

SQRT() is the square root operator

ρ is the estimated internal consistency reliability

σ_x is the standard deviation of the observed scores (in this case, Scaled Scores)

Table 10 and Table 11 summarize the distribution of CSEM values for the 2018–2019 data, overall and by grade level for the Star Math High School Algebra and Geometry tests. The overall average and grade specific CSEM on the Unified scale across all high school grades was 18 scaled score units for both Algebra and Geometry on the Unified Scale. The overall average and the grade specific CSEM based on the Unified scale is the same (18) across all grades for Algebra and Geometry. The overall average CSEM on the Enterprise scale across all grades is 30 scaled score units for Algebra and 29 for Geometry. The grade specific CSEM is 30 across all grades for Algebra. For Geometry, the grade specific CSEMs range from 29 in grades 10–12 to 30 in grade 9.

Because the standard error of measurement (SEM) is scale dependent, there are differences in the reported SEMs between the Star Math High School Algebra and Geometry Unified and Enterprise scales. Overall, the lower SEM values in Unified scale, compared to those in Enterprise scale, reflect the differences between the Unified and Enterprise scale score ranges. Neither of these is “better,” as the reliability estimates are the same for both scale.

Table 10: Algebra Standard Error of Measurement

Algebra		Standard Error of Measurement—Unified			Standard Error of Measurement—Classic		
		Conditional			Conditional		
Grade	Sample Size	Average	Standard Deviation	Global	Average	Standard Deviation	Global
9	3,000	18	1.6	18	30	2.6	29
10	4,500	18	1.5	18	30	2.4	29
11	4,000	18	1.8	18	30	3.0	29
12	1,500	18	1.4	18	30	2.3	29
Overall	13,000	18	1.6	18	30	2.6	29

Table 11: Geometry Standard Error of Measurement

Geometry		Standard Error of Measurement—Unified			Standard Error of Measurement—Classic		
		Conditional			Conditional		
Grade	Sample Size	Average	Standard Deviation	Global	Average	Standard Deviation	Global
9	4,000	18	1.7	17	30	2.7	29
10	5,000	18	1.1	17	29	1.8	29
11	1,500	18	1.1	17	29	1.8	28
12	500	18	0.9	17	29	1.5	28
Overall	11,000	18	1.3	17	29	2.2	29

Validity

Test validity was long described as the degree to which a test measures what it is intended to measure. A more current description is that a test is valid to the extent that there are evidentiary data to support specific claims as to *what* the test measures; the *interpretation* of its scores; and the uses for which it is recommended or applied. Evidence of test validity is often indirect and incremental, consisting of a variety of data that in the aggregate are consistent with the theory that the test measures the intended construct(s), or is suitable for its intended uses and interpretations of its scores. Determining the validity of a test involves the use of data and other information both internal and external to the test instrument itself.

Content Validity

One touchstone is content validity, which is the relevance of the test questions to the attributes or dimensions intended to be measured by the test. The test items should be representative of the content domain, and the content domain should be relevant to the proposed interpretation of test scores. The content of the item bank and the content balancing specifications that govern the administration of each test together form the foundation for "content validity" for the Star Math High School Algebra and Geometry assessments.

Star Math Enterprise has two options for testing high school students: One is the current Star Math Enterprise, to be known as "Integrated Math"; Star Integrated Math follows a single blueprint prescription for students in grades 9–12, with no differentiation of subject matter by grade; The other is "Star Math High School," which encompasses two distinct assessments: Algebra and Geometry. These assessments are intended to be administered to students currently enrolled in courses of the same name; students might take any one of these courses in any grade.

Star Math High School Algebra and Geometry are separate tests; both are 34-item tests administered adaptively, with the content of each separate assessment prescribed by the specifications in the blueprints.

Some noteworthy features of the blueprints for the Star Math High School assessments are:

1. Each blueprint specifies the number of items from each of several skill sets, and associated pre-requisite skills, to be administered to all students taking the assessment, regardless of their enrolled grade.

2. The content of the Star Math High School assessments' item banks are based in part on extensive analysis of existing curricula and standards. Additionally, the Star Math High School assessments' item banks are aligned to the developed Common Core of State Standards (CCSS).

3. Like the Star Math Enterprise test, Star Math High School tests are 34 items in length. This should ensure a high level of measurement precision of Star Math High School Algebra and Geometry tests and hence improve their psychometric "reliability."

4. Items are organized into skill sets similar to widely accepted standards, such as those of NCTM (National Council of Teachers of Mathematics) and CCSS. Algebra includes 24 skill sets; Geometry has 13. The number of items to be administered from each skill set is roughly proportional to the number of distinct skills subsumed under the skill set. In addition to the high school level skill sets, pre-requisite skills normally taught at earlier grades than High School Algebra and Geometry are also identified in the blueprints, and items measuring the prerequisite skills may be administered as part of the tests. These prerequisite skills items tend to be lower in calibrated difficulty than the high school algebra and geometry items and are included for the purpose of assessing students who are performing below the high school level.

5. Within each skill set, a maximum of 1 or 2 from any single skill set is prescribed; the intent of these maxima is to ensure that each test spans a range of content, rather than a concentration of multiple items measuring the same skill. The maximum numbers sum to less than 34 items; the purpose of this is to leave some degrees of freedom so that items at or near target difficulty levels can be selected in a manner less constrained by content specifications, toward the end of each student's test.

6. The item bank is large. Each of the two Star Math High School assessments includes an item bank containing several hundred test items, representing a wide range of different skill sets, as well as prerequisite skills from pre-high school grades. The content of each of the Algebra and the Geometry tests includes more than 50 discrete skills.

7. In-depth coverage of content validity are discussed in detail in the "Content and Item Development" chapter as an integral part of the test items that form the basis of the Star Math High School version.

Construct Validity

Construct validity, which is the overarching criterion for evaluating a test, investigates the extent to which a test measures the construct(s) that it claims to be assessing. Establishing evidence of construct validity involves the use of data and other information external to the test instrument itself. For example, Star Math High School Algebra and Geometry tests claims to provide an estimate of a student's Algebra or Geometry achievement level. Therefore, demonstration of Algebra or Geometry construct validity rests on the evidence that the test indeed does so.

Scores for Algebra and Geometry tests should correlate highly with other accepted measures of mathematics achievement and competence. This section deals with both internal and external evidence of the validity of Star Math High School Algebra and Geometry tests as assessments of Mathematics achievement and competence.

Internal Validity Evidence

Star Math High School Algebra and Geometry tests are applications of item response theory (IRT); each test item's difficulty has been calibrated using the Rasch 1-parameter logistic IRT model. One of the assumptions of the Rasch model is unidimensionality: that a test measures only a single construct such as mathematics achievement in the case of the Algebra or Geometry tests. One statistical method of evaluating the dimensionality of a test is by conducting factor analyses. However, there are two challenges when assessing the dimensionality of the tests by means of methods such as factor analysis:

1. Star Math High School Algebra and Geometry tests are computerized adaptive tests (CATs). Detecting dimensionality in adaptive assessments is challenging because each CAT test has somewhat different content from tests administered to other students or to the same student at other times. Different persons respond to different items, resulting in a sparse data matrix that makes conducting factor analysis more challenging for CAT data than for linear or fixed form assessments.
2. While content Domain subscores can be computed to circumvent the sparse data matrix problem, this is not possible with the Star Math High School tests. As described above, Star Math High School Algebra includes 24 skill sets; Geometry has 13 skill sets. These skill sets roll up to a single domain—either Algebra or Geometry—rendering domain-level factor analyses impossible as the minimum three variables requirement would not be met in either the Algebra tests or the Geometry tests.

Given the challenges and the limitations with the data, another source of construct validity evidence is ensuring good item fit to the scoring model. The unidimensional Rasch model is used to scale test items and retain only those items that indicate good fit to the Rasch model. If the items fit the Rasch model well, that indicates that there is only one underlying attribute, and that is the attribute we intend to measure. In addition, when Star Math High School tests scores correlate highly with other accepted procedures and measures that are used to determine math achievement, that is considered evidence of construct validity.

Relationship with Other Tests of Mathematics Achievement

When two different tests designed to measure the same or similar construct correlate strongly and positively, those correlations are viewed as evidence of construct validity.

Correlations with Florida End-of-Course (EOC) Assessments

Student performance on the Algebra and Geometry tests of the Florida EOC assessments taken during the 2017–2018 and 2018–2019 school years was compared to performance on Star Math High School Algebra and Geometry tests.

A small sample of concurrent assessments (Star tests taken within 30 days before or after the mid-date of the Florida state test window) taken by students in grades 11 and 12 was available for Algebra and suggested a strong, positive relationship ($r = 0.85$, $p < .001$). However, a concurrent sample was not available for Geometry.

Estimates of predictive validity were also computed using Star assessments taken within the same school year, but more than 30 days prior to, the Florida EOC assessments mid-date. Correlations by grade and overall for both Algebra and Geometry are reported in Table 12 and Table 13.

Table 12: Algebra Correlations—Predictive Sample

Grade	r	p	n
9	0.639	< 0.001	2,023
10	0.541	< 0.001	388
11	0.605	< 0.001	30
12			3
Overall	0.744	< 0.001	4,166

Table 13: Geometry Correlations—Predictive Sample

Grade	r	p	n
9	0.601	< 0.001	1,241
10	0.572	< 0.001	1,533
11	0.567	< 0.001	471
12	0.673	< 0.001	121
Overall	0.712	<. 0001	3,530

Correlations with Star Math Enterprise Assessments

Student performance on the Star Math High School Enterprise tests taken during the 2018–2019 school years was compared to performance on Star Math High School Algebra and Geometry tests taken the same year. Samples were first drawn from students who took the Algebra or Geometry tests at each high school grade. These samples were then matched with concurrent students who had taken the Star Math assessments.

Concurrent correlations were computed for Star Math and Algebra, Star Math and Geometry taken within the same school year. Students who took the Star Math and Algebra or the Star Math and Geometry within 30 days were used to compare the correlations. Of the concurrent sample, there were students who took both the Star Math High School subject tests and these correlations were computed. It should be noted that the time elapsed between the two subject tests may not necessarily be 30 days. The 30-day rule was imposed on the correlations between Star Math and each of the subject tests.

Table 14 presents the concurrent correlations for the Star Math Enterprise and subject tests in High School. The N represents the sample that took each of the tests in the concurrent sample. Note that the sample sizes decreased from grade 9 to grade 12.

Table 14: Concurrent Correlations—Star Math Enterprise and Algebra and Geometry

Grade	Test	Star Math Enterprise	Algebra	Geometry
9	N	1,928	10,226	3,188
	SME		0.73	0.77
	Algebra			0.66
	Geometry			1
10	N	1,209	3,017	4,629
	SME		0.78	0.74
	Algebra			0.58
	Geometry			1
11	N	750	2,510	1,268
	SME		0.83	0.67
	Algebra			0.56
	Geometry			1
12	N	241	1,114	388
	SME		0.76	0.72
	Algebra			0.64
	Geometry			1

The concurrent correlations between the Star Math Enterprise and subject tests were all substantially high For Star Math High School Algebra, the correlations ranged from 0.73 to 0.83, with an average of 0.78. For Star Math High School Geometry, the correlations ranged from 0.67 to 0.77, with an average of 0.73. Correlations between the Star Math High School Algebra and Geometry ranged from 0.56 to 0.66. These correlations provide evidence that Star Math Enterprise, Algebra and Geometry measure a similar construct of mathematical achievements to some extent but because of the content differences among the three tests, each also measures skill sets that are unique to the particular test.

Summary of Star Math High School Validity Evidence

The validity data presented in this technical manual includes evidence of Star Math High School's content and construct validity, as well as strong measures of association with math achievement levels on other independent assessments of mathematical achievements.

The first line of evidence of validity of the Star Math High School Algebra and Geometry tests is the content-related evidence from the two tests' item banks. Each of these two tests is based on a bank of test items that measure a comprehensive range of core skills that have been specified in curriculum standards for courses in high school Algebra or Geometry. Blueprints for the two tests specify the number of items measuring each of a wide range of skill sets that are to be administered within each individual test.

The correlations between Star Math High School Algebra and Geometry and other assessments, such as the Star Math Enterprise test and the independent Florida End-of-Course assessments, provide strong support for the claim that Star Math High School Algebra and Geometry are measures of mathematical achievement at high school levels. These validity results, combined with the supporting evidence of the scale score analysis as well as the reliability and minimization of SEM estimates in the "Reliability and Measurement Precision" chapter, provide a quantitative demonstration of how well the Star Math High School Algebra and Geometry perform as valid measures of high school mathematical achievement.

Norming

Star Math High School comprises two separate tests: Algebra and Geometry. This chapter describes the test score norms for each of the Star Math High School assessments.

Background

The current version of Star Math High School was introduced in June 2013 and is standards-based; it assesses a wide variety of skills and instructional standards in both Algebra and Geometry as two separate tests.

Norms for the two Star Math High School tests were first developed in 2016 for introduction at the start of the 2016–17 school year. This chapter describes the development of the 2016 norms. These norms were created for Algebra and Geometry separately.

The 2016 Star Math High School Norms

After the development of the 2016 Star Math High School, a new reporting scale was implemented, called the Unified scale. The Unified scale is a new linear transformation of the Star Math High School Rasch scores to a scale that shares features with a new scale developed for use with the general Star Math, Star Reading, and Star Early Literacy. The introduction of the Star Unified Scale provides a common scale that makes it possible for the first time to report performance on all Star assessments on the same scale.

The original Star Math High School scale is now referred to as the “Enterprise” score scale and will be available during the planned transition to the Unified scale as the default reporting scale. This chapter includes displays of normative summary data using both the Enterprise and the Unified scales.

The 2016 Star Math High School norms are based on Star Math High School Algebra and Geometry user data collected over the course of three full school years from August 2013 to June 2016. A single set of norms for all grades is reported for each of the high school tests. Students participating in the norming studies took Algebra or Geometry assessments between August 15, 2013 and June 30, 2016. Students took these Star Math High School tests under normal test administration conditions. No specific norming tests were developed and no deviations were made from the usual tests administration.

Thus, students in the norming sample took Star Math High School tests as they are administered in everyday use.

Sample Characteristics

During the norming period, a total of 65,816 US students in grades 8–12 took the Star Math High School Algebra test, and 28,969 took the Geometry test. These tests were administered using Renaissance servers hosted by Renaissance Learning. The first step in sampling was to select a representative sample of students who had tested in the fall, in the spring, or in both the fall and spring of the three school years represented in the norming studies. These norming data were a convenience sample drawn from the Star Math High School customer base. Steps to ensure the resulting norms were nationally representative of the high school grades US student population with regard to certain important characteristics were undertaken.

Because the Star Math High School tests cater to a small subset of the K–12 student body and the tests had only been used for a few years, a post-stratification procedure used to adjust the sample proportions to the approximate national proportions on three key variables: geographic region, district socio-economic status, and district/school size was explored. These three variables were chosen because they had previously been used in Star Math norming studies to weight the norming sample to approximate nationally representative samples, are known to be related to test scores, and were readily available for the schools in the Renaissance hosted database. However, a decision was made to not weight the scores as the weights were unstable due to the smaller sample sizes in the Star Math High School tests.

The final norming sample size, after selecting only students with test scores in either the fall or the spring or both fall and spring in the norming data was 43,089 for Algebra and 19,412 for Geometry students in grades 8–12. There were 34,478 assessments in the fall Algebra norming sample and 16,463 assessments in the spring Algebra norming sample; 7,144 students were included in both the fall and spring norming samples. These Algebra students came from 426 schools across 50 states and the District of Columbia.

There were 15,435 assessments in the fall Geometry norming sample and 7,318 assessments in the spring Geometry norming sample; 3,243 students were included in both norming samples. These Geometry students came from 224 schools across 50 states and the District of Columbia.

Table 15 and Table 16 provide a breakdown of the number of students participating per grade in both fall and spring for Algebra and Geometry, respectively.

Table 15: Numbers of Students per Grade in the Fall and Spring Norms Samples—Algebra

Grade	Fall	Spring
	N	N
8	8,417	6,447
9	15,127	6,302
10	5,018	2,077
11	4,520	1,348
12	1,396	289
Total	34,478	16,463

Table 16: Numbers of Students per Grade in the Fall and Spring Norms Samples—Geometry

Grade	Fall	Spring
	N	N
8	856	549
9	3,728	2,158
10	8,263	3,653
11	2,132	838
12	456	120
Total	15,435	7,318

National estimates of US student population characteristics were obtained from two entities: the National Center for Educational Statistics (NCES) and Market Data Retrieval (MDR).

- National population estimates of students' demographics (ethnicity and gender) in grades 1–12 were obtained from NCES; these estimates were from 2013–14 for private schools and 2014–15 for public schools, the most recent data available. National estimates of race/ethnicity were computed using the NCES data based on single race/ethnicity and also a multiple-race category. The NCES data reflect the most recent census data from the US census bureau.
- National estimates of school-related characteristics were obtained from May 2016 Market Data Retrieval (MDR) information. The MDR database

contains the most recent data on schools, some of which may not be reflected in the NCES data.

While factors such as region, school/district enrollment, district socio-economic status, and location are important consideration in determining how population-representative a given norming sample is, the smaller sample sizes in Algebra and Geometry resulted in a lot of missing data for meaningful sample weights analyses. As such, schools sampled in the Star Math High School norms were grouped by geographic region only for informational purposes.

Table 17 and Table 18 on page 52 show national percentages of children in grades 8–12 by geographic region, along with the corresponding percentages in the Algebra and the Geometry norming samples, respectively. MDR estimates of geographic region were based on the four broad areas identified by the National Educational Association as Northeastern, Midwestern, Southeastern, and Western regions. The specific states in each region are shown below.

Geographic Region

Using the categories established by the National Center for Education Statistics (NCES), students were grouped into four geographic regions as defined below: Northeast, Southeast, Midwest, and West.

Northeast

Connecticut, District of Columbia, Delaware, Massachusetts, Maryland, Maine, New Hampshire, New Jersey, New York, Pennsylvania, Rhode Island, Vermont

Southeast

Alabama, Arkansas, Florida, Georgia, Kentucky, Louisiana, Mississippi, North Carolina, South Carolina, Tennessee, Virginia, West Virginia

Midwest

Iowa, Illinois, Indiana, Kansas, Minnesota, Missouri, North Dakota, Nebraska, Ohio, South Dakota, Michigan, Wisconsin

West

Alaska, Arizona, California, Colorado, Hawaii, Idaho, Montana, New Mexico, Nevada, Oklahoma, Oregon, Texas, Utah, Washington, Wyoming.

Table 17: Algebra Sample Characteristics by Region: National Population Estimates and Sample Estimates

		National Estimates	Fall Norming Sample	Spring Norming Sample
Region	Midwest	20.9%	41.2%	35.4%
	Northeast	18.4%	18.0%	30.4%
	Southeast	25.1%	23.1%	16.8%
	West	35.5%	17.7%	17.3%

Table 18: Geometry Sample Characteristics by Region: National Population Estimates and Sample Estimates

		National Estimates	Fall Norming Sample	Spring Norming Sample
Region	Midwest	20.9%	39.7%	34.8%
	Northeast	18.4%	20.3%	34.5%
	Southeast	25.1%	19.0%	9.9%
	West	35.5%	21.0%	20.8%

Table 19 and Table 20 on the next page provide information on the demographic characteristics of students in the sample and national percentages provided by NCES, for Algebra and Geometry, respectively. No weighting was done on the basis of these demographic variables; they are provided to help describe the sample of students and the schools they attended. Because Star assessment users do not universally enter individual student demographic information such as gender and ethnicity/race, some students were missing demographic data, and the sample summaries in Table 19 and Table 20 are based on only those students that had gender and ethnicity information available.

Table 19: Algebra Sample Characteristics by Gender and Ethnicity: National Population Estimates and Sample Estimates

		National Estimates	Fall Norming Sample	Spring Norming Sample
Gender	Female	49.5%	43.4%	42.6%
	Male	50.5%	43.5%	43.5%
	Unknown	—	13.1%	13.9%
Ethnicity	American Indian	1%	0.8%	0.5%
	Asian	5%	3.1%	2.3%
	Black	15%	11.7%	7.4%
	Hispanic	26%	6.4%	5.7%
	White	50%	36.7%	45.5%
	Mixed Race/Ethnicity	3%	—	—
	Unknown	—	41.4%	38.6%

Table 20: Student Gender and School Information: National Estimates and Samples Percentages

		National Estimates	Fall Norming Sample	Spring Norming Sample
Gender	Female	49.5%	44.9%	46.9%
	Male	50.5%	43.8%	43.4%
	Unknown	—	11.4%	9.7%
Ethnicity	American Indian	1%	0.2%	0.3%
	Asian	5%	4.5%	3.4%
	Black	15%	10.2%	6.4%
	Hispanic	26%	5.2%	6.2%
	White	50%	36.1%	38.6%
	Mixed Race/Ethnicity	3%	—	—
	Unknown	—	43.8%	45.1%

Test Administration

All students took the current version of Star Math High School tests under normal administration procedures. Some students in the normative sample took the assessment two or more times within the norming windows; scores from their initial test administration in the fall and the last test administration in the spring were used for computing the norms.

Data Analysis

Student test records were compiled from the complete database of Star Math High School Renaissance users. Data spanned three school years from August 2013 to June 2016. Students' Rasch scores on their first Star Math Algebra or Geometry test taken during the first or the second month of the school year based on grade placement were used to compute norms for the fall; students' Rasch scores on the last Star Math Algebra or Geometry test taken during the 8th or the 9th month of the school year were used to compute norms for the spring. Interpolation was used to estimate norms for times of the year between the first month in the fall and the last month in the spring. The norms were based on the distribution of Rasch scores across all grades.

Norms were developed based on the Rasch ability estimates and then transformed to the Star Math High School Enterprise scores. Table 21 and Table 22 provide descriptive statistics for each grade with respect to the Algebra normative sample performance, in the Enterprise and Unified scaled score units, respectively. Table 23 and Table 24 provide descriptive statistics for each grade with respect to the Geometry normative sample performance, in the Enterprise and Unified scaled score units.

Table 21: Descriptive Statistics for Scaled Scores by Grade for the Algebra Norming Sample in the Enterprise Scale

Grade	Fall Enterprise Scaled Scores			Spring Enterprise Scaled Scores		
	N	Mean	Standard Deviation	N	Mean	Standard Deviation
8	8,417	827	53	6,447	875	66
9	15,127	795	73	6,302	835	72
10	5,018	819	85	2,077	855	91
11	4,520	824	73	1,348	856	79
12	1,396	804	82	289	837	76

Table 22: Descriptive Statistics for Scaled Scores by Grade for the Algebra Norming Sample in the Unified Scale

Grade	Fall Unified Scaled Scores			Spring Unified Scaled Scores		
	N	Mean	Standard Deviation	N	Mean	Standard Deviation
8	8,417	1,115	32	6,447	1,145	40
9	15,127	1,096	45	6,302	1,121	44
10	5,018	1,110	52	2,077	1,133	56
11	4,520	1,113	45	1,348	1,133	48
12	1,396	1,102	51	289	1,121	47

Table 23: Descriptive Statistics for Scaled Scores by Grade for the Geometry Norming Sample in the Enterprise Scale

Grade	Fall Enterprise Scaled Scores			Spring Enterprise Scaled Scores		
	N	Mean	Standard Deviation	N	Mean	Standard Deviation
8	856	872	52	549	960	96
9	3,728	859	50	2,158	925	78
10	8,263	814	59	3,653	868	75
11	2,132	798	67	838	841	80
12	456	797	77	120	831	71

Table 24: Descriptive Statistics for Scaled Scores by Grade for the Geometry Norming Sample in the Unified Scale

Grade	Fall Unified Scaled Scores			Spring Unified Scaled Scores		
	N	Mean	Standard Deviation	N	Mean	Standard Deviation
8	856	1,143	32	549	1,197	59
9	3,728	1,135	30	2,158	1,176	48
10	8,263	1,108	36	3,653	1,141	46
11	2,132	1,098	41	838	1,124	49
12	456	1,097	47	120	1,118	43

Score Definitions

Types of Test Scores

In a broad sense, Star Math High School Algebra and Geometry software provides three different types of test scores that measure student performance in different ways:

- *Scaled scores.* Star Math High School Algebra and Geometry assessments create a virtually unlimited number of test forms as it dynamically interacts with the students taking the test. In order to make the results of all tests comparable, and in order to provide a basis for deriving the other types of test scores described below, it is necessary to convert the results of Star tests to scores on a common scale. Star software does this in two steps. First, maximum likelihood is used to estimate each student's score on the Rasch ability scale, based on the difficulty of the items administered, and the pattern of right and wrong answers. Second, the Rasch ability scores are converted to scaled scores. Two different score scales are now available in Star assessments: the original scaled scores, which are referred to as "Enterprise" scaled scores; and a new score, expressed on the "Unified" score scale, which was introduced with the 2017–2018 school year.

  Enterprise Scale Scores

  For Star tests, the "Enterprise" scale scores are the same scores that have been reported continuously since Star Math High School Algebra and Geometry assessments were introduced in 2013. The Enterprise scaled scores range from 0 to 1400.

  Unified Scale Scores

  Renaissance developed a single score scale that applies to all Star assessments: The Unified score scale. That development began with equating each test's underlying Rasch ability scales to a common Rasch scale; the result was the "unified Rasch scale," which is an extension of the Rasch scale used in Star Reading. The next step was to develop an integer scale based on the unified Rasch scale, with scale scores anchored to important points on the original Enterprise score scales of both tests. The end result was a reported score scale that extends from 200 to 1400.

Star Math High School Algebra and Geometry can be reported on the Unified scale with a score range of 600 to 1400. One benefit of the Unified scale is an improvement in certain properties of the scale scores: scores on both tests are less variable from grade to grade; measurement error is likewise less variable; and Unified score reliability is slightly higher than that of the Enterprise scores.

- *Criterion-referenced scores* describe what a student knows or can do, relative to a specific content domain or to a standard. Such scores may be expressed either on a continuous score scale or as a classification. An example of a criterion-referenced score on a continuous scale is a percent-correct score, which expresses what proportion of test questions the student can answer correctly in the content domain. An example of a criterion-referenced classification is a proficiency category on a standards-based assessment: the student may be said to be "proficient" or not, depending on whether his score equals, exceeds, or falls below a specific criterion (the "standard") used to define "proficiency" on the standards-based test. The domain scores and mastery classification charts in the Diagnostic Report are criterion-referenced.
- *Norm-referenced scores* compare a student's test results to the results of other students who have taken the same test. In this case, scores provide a relative measure of student achievement compared to the performance of a group of students at a given time. Percentile Ranks and Grade Equivalents are the two primary norm-referenced scores provided by Star Math High School Algebra and Geometry software. Both of these scores are based on a comparison of a student's test results to the data collected during the development of the 2016 Star Math High School Algebra and Geometry norms.

Grade Equivalent (GE)

A Grade Equivalent (GE) indicates the normal grade placement of students for whom a particular score is typical. If a student receives a GE of 10.0, this means that the student scored as well on Algebra or Geometry as did the typical student at the beginning of grade 10. It does not necessarily mean that the student has mastered math objectives at a tenth-grade level, only that he or she obtained a Scaled Score as high as the average beginning tenth-grade student in the norms group.

GE scores are often misinterpreted as though they convey information about what a student knows or can do—that is, as if they were criterion-referenced scores. To the contrary, GE scores are norm-referenced.

Two sets of GE scores are used in Star Math High School; one in Algebra and the other in Geometry. The scale divides the academic year into 10 monthly increments and is expressed as a decimal with the unit denoting the grade level and the individual "months" in tenths. Because the Star Math High School tests norms are based on fall and spring score data only, monthly GE scores are derived through interpolation by fitting a curve to the month-by-month medians.

The GE scale is not an equal-interval scale. For example, an increase of 50 Scaled Score points might represent only three or four months of GE change at the lower grades, but this same increase in Scaled Scores may signify over a year of GE change in the high school grades. This occurs because student growth in math proficiency (and other academic areas) is not linear; proficiency develops much more rapidly in the lower grades than in the middle to upper grades. Algebraic manipulations of the GE scores are, therefore, not recommended.

Comparing Star Math High School GEs with Those from Conventional Tests

Because Star Math High School assessments adapt to the proficiency level of the student being tested, the GE scores that the tests provide are more consistently accurate across the achievement spectrum than those provided by conventional paper-and- pencil test instruments. In addition, Grade Equivalent scores obtained using conventional test instruments are less accurate when a student's grade placement and GE score differ markedly. It is not uncommon for a fourth- grade student to obtain a GE score of 8.9 when using a conventional test instrument. However, this does not necessarily mean that the student is performing at a level typical of an end-of-year eighth-grader. More likely, it means that the student answered all, or nearly all, of the items correctly on the conventional test and thus performed beyond the range of the fourth- grade test.

On the other hand, Star Math High School GE scores are more consistently accurate, even as a student's achievement level deviates from the level of grade placement. A student may be tested on any level of material depending on his or her actual performance on the test. Throughout a Star Math High School test, students are tested on items of an appropriate level of difficulty, based on their individual level of achievement.

Percentile Rank (PR)

Percentile Rank (PR) scores indicate the percentage of students in the same grade and at the same point of time in the school year who obtained scores lower than the score of a particular student. In other words, Percentile Ranks show how an individual student's performance compares to that of his or her same-grade peers on the national level. For example, a Percentile Rank of 85 means that the student is performing at a level that exceeds 85% of other students in that grade at the same time of the year. Percentile Ranks simply indicate how a student performed compared to others who took Star Math High School tests as a part of the national norming study. PRs range from 1–99.

The PR scale is not an equal-interval scale. For example, a grade placement of 7.7 and a Star Math High School Geometry Enterprise Scaled Score of 933 correspond to a PR of 80, and, using the same grade placement, a Scaled Score of 971corresponds to a PR of 90. Thus, a difference of 38 Scaled Score points represents a 10-point difference in PR.

However, for another student at the same grade placement, a Scaled Score of 878 corresponds to a PR of 50, and a Scaled Score of 892 corresponds to a PR of 60. While there is now only a 14-point difference in Scaled Scores, there is still a 10-point difference in PR. For this reason, PR scores should not be averaged or otherwise algebraically manipulated. NCE scores, described below, are much more appropriate for these types of calculations.

Normal Curve Equivalent (NCE)

Normal Curve Equivalents (NCEs) are scores that have been scaled in such a way that they have a normal distribution, with a mean of 50 and a standard deviation of 21.06 in the normative sample for a specific grade for a given test. Because NCEs range from 1–99, they appear similar to Percentile Ranks, but they have the advantage of being based on an equal interval scale. That is, the difference between two successive scores on the scale has the same meaning throughout the scale. Because of this feature, NCEs are useful for purposes of statistically manipulating norm-referenced test results, such as interpolating test scores, calculating averages, and computing correlation coefficients between different tests. For example, in Star Math High School score reports, average Percentile Ranks are obtained by first converting the PR values to NCE values, averaging the NCE values, and then converting the average NCE back to a PR.

Table 26 on page 62 in the Conversion Tables chapter lists the NCEs corresponding to integer PR values and facilitates the conversion of PRs to NCEs.

Grade Placement

Star Math High School software uses students' grade placement values when determining norm-referenced scores. The values of PR (Percentile Rank) and NCE (Normal Curve Equivalent) are based on what Scaled Score the student achieved; there are no grade differentiations for the high school tests. For example, a high school student in the seventh month with a Scaled Score of 902 in Geometry would have a PR of 66, while a high school student in the eighth month with the same Scaled Score would have a PR of 62.

Thus, it is crucial that student records indicate the proper grade and month within grade when students take a Star Math High School Algebra or Geometry test, and that any testing in July or August reflects the proper understanding of how Star software deals with those months in determining grade placement.

Indicating the Appropriate Grade Placement

The numeric representation of a student's grade placement is based on the specific month in which he or she takes a test. Although teachers indicate a student's grade level using whole numbers, the Star Math High School tests software automatically adds fractional increments to that grade based on the month of the test. To determine the appropriate increment, Star software considers the standard school year to run from September–June and assigns increment values of 0.0–0.9 to these months (see Table 25). The increment values for July and August depend on the school year setting:

- If teachers will use the July and August test scores to evaluate the student's math performance at the beginning of the year, in the Renaissance program, make sure the start date for that school year is before your testing in July and August. Grades are automatically increased by one level in each successive school year, so promoting students is not necessary. In this case, the increment value for July and August is 0.00 because these months are at the beginning of the school year.
- If teachers will use the test scores to evaluate the student's math performance at the end of the school year, make sure the end date for that school year falls after your testing in July and August. In this case, the increment value for July and August is 0.99 because these months are at the end of the school year that has passed.

Table 25 summarizes the increment values assigned to each month.

Table 25: Incremental Grade Placement Values per Month

Month	Decimal Increment	Month	Decimal Increment
July	0.0 or 0.99[a]	January	0.4
August	0.0 or 0.99[a]	February	0.5
September	0.0	March	0.6
October	0.1	April	0.7
November	0.2	May	0.8
December	0.3	June	0.9

a. Depends on the school year entered.

If your school follows the standard school calendar used in Star software and you will not be testing in the summer, assigning the appropriate grade placements for your students is automatic.

However, if you're going to test students in July or August, whether it is for a summer program or because your normal calendar extends into these months, grade placements become an extremely important issue.

To ensure the accurate determination of norm-referenced scores when testing in the summer, you must determine whether to include the summer months in the past school year or in the next school year. Student grade levels are automatically increased in the new school year. In most cases, you can use the above guidelines.

Instructions for specifying school years and grade assignments can be found at https://help.renaissance.com/RP and https://help2.renaissance.com/setup.

Compensating for Incorrect Grade Placements

Teachers cannot make retroactive corrections to a student's grade placement by editing the grade assignments in a student's record or by adjusting the increments for the summer months after students have tested. In other words, the Star software cannot go back in time and correct scores resulting from erroneous grade placement information. Thus, it is extremely important for the test administrator to make sure that the proper grade placement procedures are followed.

Conversion Tables

Table 26: Percentile Rank to Normal Curve Equivalent Conversions

PR	NCE	PR	NCE	PR	NCE	PR	NCE
1	1.0	26	36.5	51	50.5	76	64.9
2	6.7	27	37.1	52	51.1	77	65.6
3	10.4	28	37.7	53	51.6	78	66.3
4	13.1	29	38.3	54	52.1	79	67.0
5	15.4	30	39.0	55	52.6	80	67.7
6	17.3	31	39.6	56	53.2	81	68.5
7	18.9	32	40.1	57	53.7	82	69.3
8	20.4	33	40.7	58	54.2	83	70.1
9	21.8	34	41.3	59	54.8	84	70.9
10	23.0	35	41.9	60	55.3	85	71.8
11	24.2	36	42.5	61	55.9	86	72.8
12	25.3	37	43.0	62	56.4	87	73.7
13	26.3	38	43.6	63	57.0	88	74.7
14	27.2	39	44.1	64	57.5	89	75.8
15	28.2	40	44.7	65	58.1	90	77.0
16	29.1	41	45.2	66	58.7	91	78.2
17	29.9	42	45.8	67	59.3	92	79.6
18	30.7	43	46.3	68	59.9	93	81.1
19	31.5	44	46.8	69	60.4	94	82.7
20	32.3	45	47.4	70	61.0	95	84.6
21	33.0	46	47.9	71	61.7	96	86.9
22	33.7	47	48.4	72	62.3	97	89.6
23	34.4	48	48.9	73	62.9	98	93.3
24	35.1	49	49.5	74	63.5	99	99.0
25	35.8	50	50.0	75	64.2		

Appendix: Blueprint Skill Sets and Blueprint Skills

Table 27: Blueprint Skill Sets and Blueprint Skills—Algebra

Blueprint Skill Set	Blueprint Skill
Determine a linear equation	Determine an equation for a line parallel or perpendicular to a given graphed line
	Use a 2-variable equation to construct an input-output table
	Determine a linear equation in two variables that represents a table of values
	Determine an equation for a line given a graph
	Determine an equation of a line in slope-intercept form given the slope and y-intercept
	Determine an equation for a line given the slope of the line and a point on the line that is not the y-intercept
	Determine an equation of a line in point-slope or slope-intercept form given two points on the line
	Determine the slope-intercept form or the standard form of a linear equation
	Determine an equation for a line that goes through a given point and is parallel or perpendicular to a given line
	Represent a proportional relationship as a linear equation
	Determine an equation of a line in standard form given the slope and y-intercept
	Determine an equation of a line in standard form given two points on the line
Determine a nonlinear equation	Write a quadratic equation given its solutions
	WP: Determine a trigonometric function that represents a situation
	WP: Determine an exponential function that represents a situation such as exponential growth or decay
Evaluate an algebraic expression or function	Evaluate a 2-variable expression, with two or three operations, using integer substitution
	Evaluate a function written in function notation for a given value
Graph a 1-variable inequality	Relate a 1-variable inequality to its graph
Graph on a coordinate plane	Relate a graph to a 2-variable linear inequality
	Relate a graph to an equation of a parabola
	Relate a graph of an ellipse centered at the origin to its equation
	Determine the graph of a circle given the equation in standard form
	Determine the graph of a hyperbola given the equation in standard form
	Determine the graph of a vertically oriented parabola
	Determine the graph of a horizontally oriented parabola
	Determine the graph of a sine, cosine or tangent function

Table 27: Blueprint Skill Sets and Blueprint Skills—Algebra

Blueprint Skill Set	Blueprint Skill
Graph on a coordinate plane (continued)	Graph an ellipse
	Determine the graph of a piecewise-defined function
	Determine the component form of a vector represented on a graph
	Relate a graph to a polynomial function given in factored form
	Relate a graph to a square or cube root function
	Determine the ordered pair of a point in any quadrant
	Determine the graph of a linear equation given in slope-intercept, point-slope, or standard form
	Determine the graph of a given quadratic function
	Determine the graph of a line using given information
	Determine the graph of a 2-variable absolute value equation
	Determine the graph of the solution set of a system of linear inequalities in two variables
	Determine a 2-variable linear inequality represented by a graph
	Determine the graph of a 1-variable absolute value inequality
	WP: Answer a question using the graph of a quadratic function
	Relate a quadratic inequality in two variables to its graph
	Identify a complex number represented as a vector on a coordinate plane
Identify characteristics of a linear equation or function	Determine if lines through points with given coordinates are parallel or perpendicular
	Determine the coordinates of a point through which a line must pass in order to be parallel or perpendicular to a given line
	Determine the x- or y-intercept of a line given its graph
	WP: Interpret the meaning of the slope of a graphed line
	Determine the slope of a line given the coordinates of two points on the line
	Determine the x- or y-intercept of a line given a 1-variable equation
	WP: Determine a reasonable domain or range for a function in a given situation
	Determine the slope of a line given an equation in point-slope or slope-intercept form
	Determine if a relation is a function
	Determine if a function is linear or nonlinear
	Determine whether a graph or a table represents a linear or nonlinear function
	Determine the independent or dependent variable in a given situation
	Determine the domain or range of a function
	Determine if a table or an equation represents a direct variation, an inverse variation, or neither
	Determine the result of a change in a or c on the graph of y=ax^2 + c
	Determine the slope of a line given a table of values
	Determine the x- or y-intercept of a line given a 2-variable equation
	Determine the slope of a line given the graph of the line

Table 27: Blueprint Skill Sets and Blueprint Skills—Algebra

Blueprint Skill Set	Blueprint Skill
Identify the characteristics of a nonlinear equation or function	Identify the domain or range of a radical function
	Determine the domain and range of a graphed function
	Identify the vertex, axis of symmetry, or direction of the graph of a quadratic function
	Identify the end behavior, asymptotes, excluded values, or behavior near excluded values of a rational function
	WP: Interpret an interest rate, rate of change, initial amount, frequency of compounding and other parameters of an exponential function
	Determine if the inverse of a function is a function
	Determine the equation of the inverse of a linear, rational root, or polynomial function
	Determine the equation of a function resulting from a translation and/or scaling of a given function
	Identify the vertex form of a quadratic function and its maximum or minimum
Perform operations with logarithms	Determine the logarithmic form of an exponential equation
	Convert between a simple exponential equation and its corresponding logarithmic equation
	Evaluate a logarithm by converting it to exponential form
	Identify equivalent logarithmic expressions using the properties of logarithms
Relate a rule to a pattern	Use inductive reasoning to determine a rule
	Find a specified term of an arithmetic sequence given the first term and the common difference
	Find a specified term of an arithmetic sequence
	Find a specified term of an arithmetic sequence given the formula for the nth term
	WP: Solve a problem that can be represented by an arithmetic sequence
	Find a specified term of a geometric sequence given the first three terms of the sequence
	Determine the explicit formula for an arithmetic sequence
	Identify a given sequence as arithmetic, geometric, or neither
	Find a specified term of a binomial expression raised to a positive integer power
	WP: Solve a problem that can be represented by a geometric sequence
	WP: Solve a problem that can be represented by a finite geometric series
	Determine the algebraic equation that describes a pattern represented by data in a table
Simplify an algebraic expression	Determine the composition of two functions
	Represent an algebraic radical expression in exponential form
	Simplify an expression with rational exponents
	Factor a polynomial using long division
	Factor a polynomial by grouping
	Factor a difference of squares
	Factor the sum or difference of 2 cubes
	Factor a polynomial into a binomial and trinomial

Table 27: Blueprint Skill Sets and Blueprint Skills—Algebra

Blueprint Skill Set	Blueprint Skill
Simplify an algebraic expression (continued)	Simplify a monomial algebraic expression that includes fractional exponents and/or nth roots
	Multiply or divide functions
	Identify an equivalent form of a third or fourth degree polynomial, given its factored form
	Simplify an algebraic expression by combining like terms
	Multiply two monomial algebraic expressions
	Add or subtract polynomial expressions
	Simplify a rational expression involving polynomial terms
	Multiply rational expressions
	Divide a polynomial expression by a monomial
	Apply the product of powers property to a monomial algebraic expression
	Apply the power of a power property to a monomial algebraic expression
	Apply the power of a product property to a monomial algebraic expression
	Apply the quotient of powers property to monomial algebraic expressions
	Apply the power of a quotient property to monomial algebraic expressions
	Multiply two binomials of the form (ax +/- b)(cx +/- d)
	Factor the GCF from a polynomial expression
	Factor trinomials that result in factors of the form (ax +/- b)(cx +/- d)
	Simplify a monomial algebraic radical expression
	Apply terminology related to polynomials
	Multiply two binomials of the form (x +/- a)(x +/- b)
	Simplify a polynomial expression by combining like terms
	Multiply a polynomial by a monomial
	Multiply two binomials of the form (ax +/- by)(cx +/- dy)
	Multiply a trinomial by a binomial
	Factor trinomials that result in factors of the form (x +/- a)(x +/- b)
	Factor a trinomial that results in factors of the form (ax +/- by)(cx +/- dy)
	Factor the difference of two squares
	Factor a perfect-square trinomial
	Multiply monomial algebraic radical expressions
	Divide monomial algebraic radical expressions
	Divide a polynomial expression by a binomial
	Add or subtract two rational expressions with like denominators
	Add or subtract two rational expressions with unlike monomial denominators
	Apply properties of exponents to monomial algebraic expressions
	Factor a polynomial that has a GCF and two linear binomial factors
	Rationalize the denominator of an algebraic radical expression
	Add or subtract algebraic radical expressions

Table 27: Blueprint Skill Sets and Blueprint Skills—Algebra

Blueprint Skill Set	Blueprint Skill
Solve a linear equation	Solve a 2-step linear equation involving integers
	Solve a 2-step equation involving rational numbers
	Determine a solution to a 2-variable linear equation
	Solve a 1-variable linear equation with the variable on both sides
	Rewrite an equation to solve for a specified variable
	Solve a 1-variable linear equation that requires simplification and has the variable on one side
	Solve a direct or inverse variation problem
Solve a linear inequality	Solve a 2-step linear inequality in one variable
	Solve a 1-variable linear inequality with the variable on both sides
	Solve a 1-variable linear inequality with the variable on one side
	Solve a 1-variable compound inequality
	Solve a 2-variable linear inequality for the dependent variable
	Determine if an ordered pair is a solution to a 2-variable linear inequality
Solve a nonlinear equation	Solve a quadratic equation using the square root rule
	Solve a quadratic equation by factoring
	Determine an equation of a circle
	Determine the radius, center, or diameter of a circle given an equation
	Solve a quadratic equation with complex solutions
	Solve a logarithmic equation
	Solve a cubic equation
	Write the equation of a circle given its center and radius
	Determine the term needed to complete the square in a quadratic equation
	Solve a radical equation that leads to a quadratic equation
	Solve a rational equation involving terms with monomial denominators
	Solve a rational equation involving terms with polynomial denominators
	Solve a 1-variable absolute value inequality
	Solve a quadratic equation using the quadratic formula
	Solve a radical equation that leads to a linear equation
	Solve a quadratic equation by taking the square root
	Determine the solution(s) of an equation given in factored form
	Use the discriminant to determine the number of real solutions
	Solve a 1-variable absolute value equation
Solve a system of linear equations	Solve a system of three equations
	Solve a system of linear equations in two variables using any method
	Solve a number problem that can be represented by a linear system of equations
	Determine the number of solutions to a system of linear equations

Table 27: Blueprint Skill Sets and Blueprint Skills—Algebra

Blueprint Skill Set	Blueprint Skill
Determine a missing measure or dimension of a shape	Determine a missing angle measure in a triangle
Solve a problem involving the perimeter of a shape	Solve a problem involving the circumference of a circle
Determine a measurement	Determine a sine, cosine, or tangent ratio in a right triangle
	Convert between degree measure and radian measure
	Determine the value of an inverse sine, cosine, or tangent expression
Add and subtract fractions with unlike denominators	Add fractions with unlike 1-digit denominators
	Add mixed numbers with unlike denominators
Convert between an improper fraction and a mixed number	Convert an improper fraction to a mixed number
Determine a square root	Evaluate the positive square root of a perfect square
	Determine the square root of a perfect-square fraction or decimal
Multiply and divide with fractions	Determine the reciprocal of a positive whole number, a proper fraction, or an improper fraction
	Multiply a fraction by a fraction
Perform operations with integers	Add integers
	Subtract integers
	Multiply integers
	Divide integers
Perform operations with matrices, vectors, and complex numbers	Multiply a matrix by a scalar
	Add or subtract matrices
	Multiply matrices
	Solve a problem involving matrices
	Add or subtract complex numbers
	Simplify an expression involving a complex denominator
	Write an imaginary number in standard form
	Represent a system of linear equations as a single matrix equation
	Determine the inverse of a matrix
	Multiply complex numbers
	Add or subtract vectors component-wise
	Evaluate a linear combination of vectors
	Determine the determinant of a matrix
Solve a proportion, rate, or ratio	Determine if ratios are equivalent

Table 28: Blueprint Skill Sets and Blueprint Skills—Geometry

Blueprint Skill Set	Blueprint Skill
Determine a linear equation	Use a 2-variable equation to construct an input-output table
	Use a 2-variable equation to represent a relationship expressed in a table
	Determine a linear equation in two variables that represents a table of values
	Determine an equation for a line given a graph
	Determine an equation of a line in slope-intercept form given the slope and y-intercept
	Determine an equation for a line given the slope of the line and a point on the line that is not the y-intercept
	Determine an equation of a line in point-slope or slope-intercept form given two points on the line
	Determine an equation of a line in standard form given the slope and y-intercept
	Determine an equation of a line in standard form given two points on the line
Graph on a coordinate plane	Determine the graph of a circle given the equation in standard form
	Determine the graph of a line using given information
Identify characteristics of a linear equation or function	Determine the x- or y-intercept of a line given its graph
	WP: Interpret the meaning of the slope of a graphed line
	Determine the slope of a line given the coordinates of two points on the line
	Determine the x- or y-intercept of a line given a 1-variable equation
	Determine the slope of a line given an equation in point-slope or slope-intercept form
	Determine the x- or y-intercept of a line given a 2-variable equation
	Determine the slope of a line given the graph of the line
Solve a linear inequality	Solve a 2-step linear inequality in one variable
Solve a nonlinear equation	Determine an equation of a circle
	Determine the radius, center, or diameter of a circle given an equation
Determine a missing measure or dimension of a shape	Use the Pythagorean theorem to determine a length
	Determine the midpoint of a line segment given the coordinates of the endpoints
	Determine the measure of an angle formed by parallel lines and one or more transversals given an angle measure
	Determine the measure of an angle or the sum of the angles in a polygon
	Determine a length using parallel lines and proportional parts
	Determine a length using the properties of a 45-45-90 degree triangle or a 30-60-90 degree triangle
	Solve a problem involving the length of an arc
	Determine the length of a line segment, the measure of an angle, or the measure of an arc using a tangent to a circle
	Determine the measure of an arc or an angle using the relationship between an inscribed angle and its intercepted arc
	Solve a problem involving the distance formula

Table 28: Blueprint Skill Sets and Blueprint Skills—Geometry

Blueprint Skill Set	Blueprint Skill
Determine a missing measure or dimension of a shape (continued)	Solve a problem using inequalities in a triangle
	Solve for the length of a side of a triangle using the Pythagorean theorem
	WP: Determine a length or an angle measure using triangle relationships
	Determine the length of a side or the measure of an angle in congruent triangles
	Determine the length of a side in one of two similar polygons
	Determine the length of a side or the measure of an angle in similar triangles
	Determine a length given the perimeters of similar triangles or the lengths of corresponding interior line segments
	Determine a length in a triangle using a midsegment
	Determine the measure of an arc or a central angle using the relationship between the arc and the central angle
	Solve a problem involving the midpoint formula
	Determine a length or an angle measure using the segment addition postulate or the angle addition postulate
	Solve a problem involving a bisected angle or a bisected segment
	Determine the measure of an angle in a figure involving parallel and/or perpendicular lines
	Determine the measure of an angle using angle relationships and the sum of the interior angles in a triangle
	Determine a length in a triangle using a median
	Solve a problem involving a point on the bisector of an angle
	Determine a length or an angle measure using general properties of parallelograms
	Determine a length or an angle measure using properties of squares, rectangles, or rhombi
	Determine a length or an angle measure using properties of kites
	Determine a length or an angle measure using properties of trapezoids
	Determine the effect of a change in dimensions on the perimeter or area of a shape
	Determine the distance between two points on a coordinate plane
	Determine the measure of an angle formed by parallel lines and one or more transversals given algebraic expressions
	Use triangle inequalities to determine a possible side length given the length of two sides
	Determine the measure of an angle or an arc using a tangent to a circle
	Determine a missing angle measure in a triangle
	Determine a missing dimension given two similar shapes
Identify a geometric construction	Identify a geometric construction given an illustration
Identify congruence and similarity of geometric shapes	Identify a triangle congruence postulate that justifies a congruence statement
	Identify a triangle similarity postulate that justifies a similarity statement
	Identify similar triangles using triangle similarity postulates or theorems
	Identify congruent triangles using triangle congruence postulates or theorems
	Determine the coordinates of a preimage or an image given a reflection across a horizontal line, a vertical line, the line y = x, or the line y = -x
	Determine the coordinates of the image of a figure after two transformations of the same type

Table 28: Blueprint Skill Sets and Blueprint Skills—Geometry

Blueprint Skill Set	Blueprint Skill
Solve a problem involving the area of a shape	Determine a length given the area of a parallelogram
	Determine the area of a sector of a circle
	Determine the length of the radius or the diameter of a circle given the area of a sector
	Determine the measure of an arc or an angle given the area of a sector of a circle
	Determine the area or circumference of a circle given an equation of the circle
	Determine a length given the area of a kite or rhombus
	Determine a length given the area of a trapezoid
	Solve a problem involving density and area
Solve a problem involving the surface area or volume of a solid	Determine a length given the surface area of a right cylinder or a right prism that has a rectangle or a right triangle as a base
	Solve a problem involving the volume of a right pyramid or a right cone
	Determine the surface area of a sphere
	Determine the volume of a sphere or hemisphere
	Solve a problem involving the surface areas of similar solid figures
Use the vocabulary of geometry and measurement	Relate the coordinates of a preimage or an image to a translation described using mapping notation
	Relate the coordinates of a preimage or an image to a dilation centered at the origin
	Identify a relationship between points, lines, and/or planes
	Identify angle relationships formed by multiple lines and transversals
	Identify parallel lines using angle relationships
	Determine the angle of rotational symmetry of a figure
	Use deductive reasoning to draw a valid conclusion from conditional statements
	Identify a statement or an example that disproves a conjecture
	Identify a valid biconditional statement
	Determine the number of faces, edges, or vertices in a 3-dimensional figure
	Identify a cross section of a 3-dimensional shape
	Relate a net to a 3-dimensional shape
	Identify the converse, inverse, or contrapositive of a statement
Determine a measurement	Determine a sine, cosine, or tangent ratio in a right triangle
	Identify angle relationships formed by parallel lines cut by a transversal
Solve a proportion, rate, or ratio	Determine if ratios are equivalent
	WP: Determine a measure of length, weight or mass, or capacity or volume using proportional relationships

Index

G

H

I

K

M

N

P

R

S

T

U

V

About Renaissance

Renaissance® transforms data about how students learn into instruments of empowerment for classroom teachers, enabling them to guide all students to achieve their full potentials. Through smart, data-driven educational technology solutions that amplify teachers' effectiveness, Renaissance helps teachers teach better, students learn better, and school administrators lead better. By supporting teachers in the classroom but not supplanting them, Renaissance solutions deliver tight learning feedback loops: between teachers and students, between assessment of skills mastery and the resources to spur progress, and between each student's current skills and future academic growth.

RENAISSANCE®

©Copyright 2021 Renaissance Learning, Inc. All rights reserved. | (800) 338-4204 | www.renaissance.com

All logos, designs, and brand names for Renaissance's products and services, including, but not limited to, Accelerated Math, Accelerated Reader, Accelerated Reader 360, AccelScan, English in a Flash, MathFacts in a Flash, myON, myON Reader, myON News, Renaissance, Renaissance Flow 360, Renaissance Growth Platform, Renaissance Growth Alliance, Renaissance Learning, Renaissance-U, Renaissance Smart Start, Star, Star 360, Star Custom, Star Early Literacy, Star Early Literacy Spanish, Star Math, Star Math Spanish, Star Reading, Star Reading Spanish and Star Spanish, are trademarks of Renaissance. All other product and company names should be considered the property of their respective companies and organizations.

210302