STAR™ Math™

Technical Manual

Renaissance Learning
P. O. Box 8036
Wisconsin Rapids, WI 54495-8036

Toll-Free Phone: (800) 338-4204
Local: (715) 424-3636
Outside the US: 1.715.424.3636
Fax: (715) 424-4242

Email (general questions): answers@renlearn.com
Email (technical questions): support@renlearn.com
Website: www.renlearn.com

Copyright Notice

STAR Math has been reviewed for scientific rigor by the National Center on Student Progress Monitoring. It was found to meet the Center's criteria for scientifically based progress monitoring tools, including its reliability and validity as an assessment. For more details, visit http://www.studentprogress.org.

9/2012 SMRP

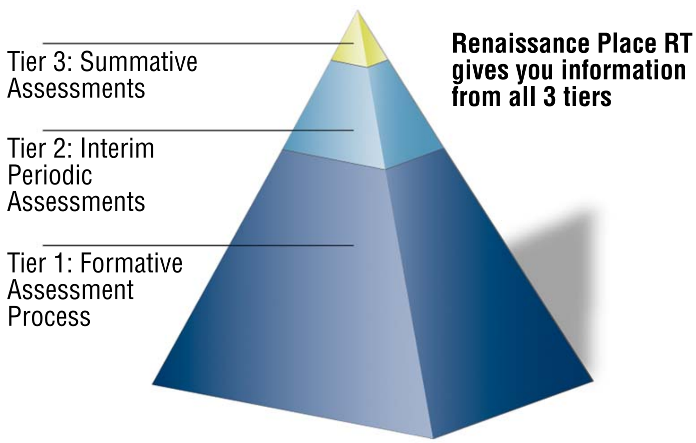
Contents

Introduction

STAR Math: Progress Monitoring Assessment

The Renaissance Place Real Time (RT) edition of STAR Math computer-adaptive test and database helps teachers accurately assess students' mathematical abilities in an average of 20 minutes.[1] This computer program also helps educators accelerate learning and increase motivation by providing immediate, individualized feedback on student academic tasks and classroom achievement. All key decision-makers throughout the district can easily access this information.

The Renaissance Place RT database stores all three levels of student information, including the Tier 2 data from STAR Math.



Tier 1: Formative Assessment Process

Formative assessment provides daily, even hourly, feedback on students' task completion, performance, and time on task. Renaissance Learning Tier 1 programs include Accelerated Reader, MathFacts in a Flash, Accelerated Math, English in a Flash, and NEO/NEO 2.

Tier 2: Interim Periodic Assessments

Interim periodic assessments help educators match the level of instruction and materials to the ability of each student, measure growth throughout the year, predict outcomes on mandated state tests, and track growth in student achievement longitudinally, facilitating the kind of growth analysis recommended by state and federal organizations. Renaissance Learning Tier 2 programs include STAR Early Literacy, STAR Math, and STAR Reading.

1. Some students may require an additional 10 to 15 minutes.

Tier 3: Summative Assessments

Summative assessments provide quantitative and qualitative data in the form of state-mandated tests. The best way to ensure success on Tier 3 assessments is to monitor progress and adjust instructional methods and practice activities throughout the year using Tier 1 and Tier 2 assessments.

STAR Math Purpose

As a periodic progress monitoring system, STAR Math software serves two primary purposes. First, it provides educators with quick and accurate estimates of students' instructional math levels relative to national norms. Second, it provides the means for tracking growth in a consistent manner over long time periods for all students. This is especially helpful to school- and district-level administrators.

While the STAR Math test provides accurate normed data like traditional norm-referenced tests, it is not intended to be used as a "high-stakes" test. Generally, states are required to use high-stakes tests 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 is not intended for these purposes. Rather, because of the high correlation between the STAR Math test and high-stakes instruments, classroom teachers can use STAR Math scores to fine-tune instruction while there is still time to improve performance before the regular testing cycle. At the same time, school- and district-level administrators can use STAR Math to predict performance on high-stakes tests. Furthermore, STAR Math results can easily be disaggregated to identify and address the needs of various groups of students.

STAR Math's unique powers of flexibility and repeatability provide specific advantages for various groups:

- For students, STAR Math software provides a challenging, interactive, and brief test that builds confidence in their math ability.
- For teachers, STAR Math software facilitates individualized instruction by identifying students' current developmental levels and areas for growth.
- For principals, STAR Math software provides regular, accurate reports on performance at the class, grade, building, and district level, as well as year-to-year comparisons.
- For district administrators and assessment specialists, STAR Math software furnishes a wealth of reliable and timely data on math growth at each school and throughout the district. It also provides a valid basis for comparing data across schools, grades, and special student populations.

This manual documents the suitability of the STAR Math progress monitoring system for these purposes and presents evidence of its reliability, validity, and merits as a psychometric instrument.

STAR Math Enterprise

STAR Math Enterprise is the same as STAR Math, but with some enhanced features, including additional reports and expanded benchmark management.

In this manual, information that refers to Enterprise-only program functions will have the ENTERPRISE indicator next to them.

Scale and the Development of STAR Math Enterprise

Development of STAR Math Enterprise began with thorough analyses of the national and state-level standards, including the Common Core state standards. Once the content had been catalogued, Renaissance Learning's standards experts did the following:

- Developed Core Progress learning progression for math, including identifying sequences of knowledge and skills and of prerequisite knowledge and skills.
- Wrote thousands of test questions, each one keyed to a specific grade level and designed to measure one element of knowledge or skill specific to that level.

The next step was to place the test questions on a single scale of difficulty spanning kindergarten to high school by administering each question to a thousand or more students at appropriate grade levels. Rigorous psychometric analyses resulted in accurate placement of each test question on the STAR Math scale. Every question's difficulty is calibrated on the same scale used to report STAR Scaled Scores.

Examination of the item calibration results confirmed that the rank order of the difficulty of the STAR test items corresponded very closely to the rank order of the skills in the learning progression for math. As a result, a student's Scaled Score (representing his or her location on the STAR scale) can be mapped to the learning progression for math, enabling research-based inferences about which skills that student has likely already developed, which are ready to be developed, and which will likely develop soon.

In this way, the score from a STAR Math Enterprise test provides reliable information, not just about the skills directly related to the test questions the student actually answered correctly or incorrectly, but also about the student's degree of proficiency on the entire array of skills in the learning progression for math. STAR Math's learning continuum is research-based, robust, and supported by experts in the field of mathematics.

Design of STAR Math

One of the fundamental decisions when designing STAR Math involved the choice of how to administer the test. Because of the numerous advantages offered by computer-administered tests, it was decided to develop STAR Math as a computer software product.

The primary advantage of using computer software to administer the STAR Math test is the ability to tailor each student's test based on his or her specific responses to previous items. Paper-and-pencil tests are obviously far different from this: every student must respond to the same items in the same sequence. Using computer-adaptive procedures, however, it is possible for students to be tested using items that appropriately match their current level of proficiency. Adaptive Branching, the item selection procedure used in the STAR Math test, effectively customizes every test to the student's current achievement level.

Adaptive Branching offers significant advantages in terms of test reliability, testing time, and student motivation. First, reliability improves over paper-and-pencil tests because the test difficulty matches each individual's performance level; students do not have to fit a "one test fits all" model. With a computer-adaptive test, most of the test items to which students respond are at levels of difficulty that closely match their achievement levels. Testing time decreases because, unlike in paper-and-pencil tests, students need not be exposed to a broad range of material, some of which is inappropriate because it is either too easy for high achievers or too difficult for those with low current levels of performance. Finally, computer-adaptive assessments improve student motivation simply because of the aforementioned issues: test time is minimized and test content is neither too difficult nor too easy. Not surprisingly, most students enjoy taking STAR Math tests, and many report that it increases their confidence in math.

Another fundamental STAR Math design decision involved the format of the test items. The items had to be easily administered and objectively scored by a computer and also provide the breadth of construct coverage necessary for an assessment of math achievement. The traditional four-item multiple-choice format was chosen, based on considerations of efficiency of assessment, objectivity, and simplicity of scoring.

This manual describes two distinct versions of STAR Math assessments: STAR Math and STAR Math Enterprise. STAR Math, the original assessment, is a 24-item measure of general achievement in math, and is based on a bank of more than 2,000 test items spanning more than 200 objectives. STAR Math Enterprise is a 34-item standards-based assessment that draws its items from a bank of more than 4,000 test items measuring more than 550 skills. The Enterprise version also differs from the original version in the organization of its content, as will be noted below.

STAR Math: A fundamental design decision involved determining the organization of the content in STAR Math. Because of the great amount of overlap in content in the math construct, it is difficult to create distinct categories or "strands" for a mathematics achievement instrument. After reviewing the STAR Math test's content, curricular materials, and similar math achievement instruments, the following eight strands were identified and included in STAR Math: Numeration Concepts, Computation Processes, Word Problems, Estimation, Data Analysis and Statistics, Geometry, Measurement, and Algebra.

The STAR Math test is further divided into two parts. The first part of the test, the first sixteen items, includes items only from the Numeration Concepts and the Computation Processes strands. The first eight test items (items 1–8) are from the Numeration Concepts strand, and the following eight test items (items 9–16) are from the Computation Processes strand.

The second part of the test, or the final eight items, includes items from all of the remaining strands. Hence, items 17–24 are drawn from the following six strands: Word Problems, Estimation, Data Analysis and Statistics, Geometry, Measurement, and Algebra. The specific makeup of the strands used in the final eight items depends on the student's grade level. For example, a student in grade 1 will not receive items from the Estimation strand, but items from this strand could be administered to a student in grade 12.

The decision to weight the test heavily toward Numeration Concepts and Computation Processes resulted from the fact that these strands are fundamental to all others, and they include the content about which teachers desire the most information. Although this approach emphasizes the two strands in the first part of the test, it provides adequate content balance to assure valid assessment. Additionally, factor analysis of the various content strands supports the fundamental unidimensionality of the construct being measured in the STAR Math test.

STAR Math ENTERPRISE: The organization of the content in STAR Math Enterprise differs from that of the original STAR Math. The Enterprise version's content organization reflects current thinking, as embodied in many different sets of national and local curriculum standards. The following four domains were identified and included in STAR Math Enterprise: Numbers and Operations, Algebra, Geometry and Measurement and Data Analysis, Statistics and Probability. Within each of these domains, skills are organized into skill sets; there are 54 skill sets in all, comprising a total of over 550 core skills.

The STAR Math Enterprise test is a 34-item standards-based version of STAR Math, administered as 6 blocks of items in a single section. Each block of items contains a blend of items from the 4 domains. The number of items administered in a block varies by grade band. The item sequencing calls for more content balance at the beginning, middle, and end of the test by "spiraling" the content throughout the test, thus ensuring that the ability estimate at any point during a test is based on a broad range of content, rather than on a limited sample of skills.

Each STAR Math item was developed in association with a very specific content objective (described in "Content and Item Development" on page 14). In addition, the calibration trials included items that were expressed differently in textbooks and other reference materials, and only the item formats that provided the best psychometric properties were retained in the final item bank. For example, many questions were crafted both with and without graphics supporting the text of the question. For items containing text in either the question stem or the response choices, great care was taken to keep the text simple and the reading level as low as practical. This is particularly important with computer-adaptive testing because high-performing, lower-grade students may receive higher grade-level questions.

In an attempt to minimize the administration of inappropriate items to students, each item in the item bank is assigned a curricular placement value corresponding to the lowest grade where instruction for this content would occur. During STAR Math testing, students receive items with a maximum curricular placement value of three grades higher than their current grade level. Although this constraint does not limit the attainable scores in any way, since very difficult items still exist in the item bank within these constraints, it does help to minimize presentation of items for which the student has not yet had any formal instruction.

ENTERPRISE STAR Math Enterprise is a standards-based test that uses items that measure standards appropriate to a student's grade, or standards the student should have mastered at lower grades. It is not designed to test standards that are far above a student's actual grade level.

Test Interface

The STAR Math test interface was designed to be both simple and effective. Students can use either the keyboard or the mouse to input answers.

- If using the keyboard, students press one of the four letter keys (**A**, **B**, **C**, and **D**) and the **Enter** key (or the **return** key on Macintosh computers).
- If using the mouse, students click the answer of choice and click **Next** to complete the item response.

Practice Session

The practice session before the STAR Math test allows students to become comfortable with the test interface and to make sure that they know how to operate the software properly. Students can pass the practice session and proceed to the actual STAR Math test by answering two out of the three practice questions correctly. If a student does not do this, the program presents three more questions, and the student can pass the practice session by answering two of those three questions correctly. If the student does not pass after the second attempt, the student will not proceed to the actual STAR Math test. Students who have successfully passed a practice session within the last 180 days will not get practice questions.

Even students with low math and reading skills should be able to answer the practice questions correctly. However, STAR Math will halt the testing session and tell the student to ask the teacher for help if the student does not pass 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 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 test.

Adaptive Branching/Test Length

STAR Math's item selection branching algorithm uses a proprietary approach somewhat more complex than the simple Rasch Maximum Information IRT model. The approach used in the STAR Math test was designed to yield reliable test results by adjusting item difficulty to the responses of the individual being tested while striving to minimize test length and student frustration.

As an added measure to minimize student frustration, the first administration of the test begins with items that have a difficulty level below what a typical student at a given grade can handle, usually one or two grades below grade level.

Teachers can override the use of grade placement for determining starting difficulty by entering the current level of mathematics instruction for the student using the MIL

(Math Instruction Level). When an MIL is provided, the program uses that value to raise or lower the starting difficulty of the first test. On the second and subsequent administrations, the test begins about one grade lower than the ability last demonstrated within 180 days.

Once the testing session is underway, STAR Math software administers 24 items of varying difficulty, adapting the difficulty level of the items dynamically according to the student's responses. The average length of time required to complete a STAR Math test is between 11 and 12 minutes, with a standard deviation of about four minutes. It should be noted that unlike traditional tests, the time required for completion increases with ability. For example, students performing at and above the 90th percentile will on average require about 13 minutes to complete the test, while students performing at or below the 10th percentile require only 10 minutes.

ENTERPRISE The STAR Math Enterprise test administers 34 operational items. Practice items are bypassed if the student has passed the practice within the last 180 days. The average length of time required to complete these 34 questions is approximately 20 minutes. As with the STAR Math test, time required for completion increases with ability.

Test Repetition

Repeated testing allows teachers to measure a student's math growth over time. STAR Math can be used for multiple purposes such as screening, placement, diagnostic assessment, benchmark assessment, and outcomes measurement. It may be used weekly in progress-monitoring programs, and has been found to meet the standards of the National Center for Student Progress Monitoring for monthly assessments.

The STAR Math item bank contains more than 2,000 items created from eight different content strands. Because the STAR Math software keeps track of the specific items presented to each student from test session to test session, it does not present the same item more than once in any 75-day period. By doing so, the software keeps item reuse to a minimum. In addition, if a student is progressing in mathematics development throughout the year and from year to year, item exposure should not be an issue at all.

ENTERPRISE The STAR Math Enterprise item bank includes items measuring over 550 skills in 54 skill sets from four domains. STAR Math software keeps track of the specific items presented to each student from test session to test session, and does not present the same item more than once in any 75-day period.

More information on the content of the STAR Math item bank is available in "Content and Item Development" on page 14.

Item Time Limits

The STAR Math test has a fixed three-minute time limit for individual test items (both operational and calibration) and a fixed ninety-second time limit for practice items. A fixed time limit was chosen to avoid the complexity and confusion associated with a variable time-out period. Three minutes was chosen on the basis of calibration and norming timing data and general content testing experience.[2]

When a student has only 15 seconds remaining for a given item, a picture of a clock appears in the upper-right corner of the screen, 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 and hasn't yet pressed **Enter** or **return** (or clicked **Next**) before the item times out. In that case, the answer is accepted as correct.

The items were crafted with one minute as the maximum amount of time that a student who knew how to do the mathematics would require to complete the solution and respond. During STAR Math norming, the mean item response time was 27 seconds with a standard deviation of 25 seconds. The median was 19 seconds, and nearly all (99.7%) item responses were made within the three-minute time limit. Mean and median response times were similar at all grades. Although the incidence of maximum time limits was somewhat higher at the lowest three grades than in other grades, fewer than half of one percent of item responses reached the time limit. This was true even for first-grade students. This suggests that the time limits used for STAR Math allow ample time for nearly all students to complete the questions.

Time Limits and the STAR Math Diagnostic Report

The STAR Math Diagnostic Report includes a conditional text section in the case that a student completes the test in much less time than expected. There are two parts of the test considered in the report explanation.

The first part includes the first 16 items that appear in the test. If the student completes the first part in 107 seconds or less, the following text appears in the report:

> Time for First Part: # seconds   Time for Second Part: # seconds
>
> The time required to complete the first part of the test was very low. It may be that (Name) can do math very quickly, or that (Name) did not try very hard on the first part of the test. If you suspect the latter to be true, you may want to discuss the situation with the student and retest.

The second part includes the last 8 items that appear in the test. If the student completes the second part in 49 seconds or less, the following text appears in the report:

> Time for First Part: # seconds   Time for Second Part: # seconds
>
> The time required to complete the second part of the test was very low. It may be that (Name) can do math very quickly, or that (Name) did not try very hard on the second part of the test. If you suspect the latter to be true, you may want to discuss the situation with the student and retest.

If the student completes both parts of the test within the respective time frames, the following text appears in the report:

> Time for First Part: # seconds   Time for Second Part: # seconds
>
> The times required to complete both parts of the test were very low. It may be that (Name) can do math very quickly, or that (Name) did not try very hard on the test. If you suspect the latter to be true, you may want to discuss the situation with the student and retest.

2. After July 2009, teachers and test monitors gained the ability to extend time limits for questions for students who have special needs. The standard time limits are 90 seconds for practice questions and 180 seconds for actual test questions; the extended time limits allow 180 seconds for practice questions and 360 seconds for actual test questions.

ENTERPRISE Although the STAR Math Enterprise test is not in two parts like the STAR Math test is, similar messages will appear in the Diagnostic Report if the time taken to complete the test is considerably less than expected.

Test Security

STAR Math software includes a variety of features intended to provide adequate security to protect the content of the test and to maintain the confidentiality of the test results.

Split Application Model

In the STAR Math Renaissance Place (RP) software, when students log in, they do not have access to the same functions that teachers, administrators, and other personnel can access. Students are allowed to test, but they have no other tasks available in STAR Math RP; therefore, they have no access to confidential information. When teachers and administrators log in, they can manage student and class information, set preferences, register students for testing, and create informative reports about student test performance.

Individualized Tests

Using Adaptive Branching, every STAR Math 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 data or access or change it with other software.

Access Levels and Capabilities

Each user's level of access to a Renaissance Place program depends on the primary position assigned to that user and the capabilities the user has been granted in Renaissance Place. Each primary position is part of a user group. There are seven user groups: district administrator, district staff, school administrator, school staff, teacher, parent, and student. By default, each user group is granted a specific set of capabilities. Each capability corresponds to one or more tasks that can be performed in the program. The capabilities in these sets can be changed; capabilities can also be granted or removed on an individual level. Since users can be assigned to the district and/or one or more schools (and be assigned different primary positions at the different locations), and since the capabilities granted to a user can be customized, there are many, varied levels of access an individual user can have.

Renaissance Place RT also allows you to restrict students' access to certain computers. This prevents students from taking STAR Math tests from unauthorized computers (such as home computers). For more information on student access security, see the *Renaissance Place Real Time Software Manual*.

The security of the STAR Math 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 primary position and the capabilities that they have been granted. Personnel who log in to Renaissance Place RT (teachers, administrators, and staff) must enter a user name and password before they can access the data and create reports. Parents must also log in with a user name and password before they can access the Parent Report. Without an appropriate user name and password, personnel and parents cannot use the STAR Math RP software.

Test Monitoring/Password Entry

Test monitoring is another useful STAR Math security feature. Test monitoring is implemented using the Testing Password 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 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 user 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.

They should also ensure that scratch paper used in the testing process is gathered and discarded after each testing session. Taking these simple precautionary steps will help maintain STAR Math's security and the quality and validity of its scores.

Psychometric Characteristics

The following sections provide an overview of the content of the STAR Math test, its length in both number of items and administration time, and also its Adaptive Branching feature, the test scores it yields, and how those scores are distributed. Some of these features differ between STAR Math and STAR Math Enterprise, as noted below.

Content

Every STAR Math assessment consists of items that tap knowledge and skills from as many as eight different mathematical strands. The items comprise several sets of skills for each strand, with 17 different clusters of skills in all (9 Computation clusters and 8 Numeration clusters). The STAR Math test has 24 questions.

ENTERPRISE Every STAR Math Enterprise assessment consists of items that tap knowledge and skills from as many as four different standards-based mathematical domains. The items comprise several skill sets for each domain, with 54 skill sets in all. The STAR Math Enterprise test has 34 questions.

Content balancing specifications ensure that a specific number of items from each domain are administered in every test.

- "Appendix A: Objectives and STAR Math Items" on page 135 contains a detailed list of the skills assessed by STAR Math.
- "Appendix B: Objectives and STAR Math Enterprise Items" on page 143 contains a detailed list of the skills assessed by STAR Math Enterprise.

STAR Math Enterprise and the Common Core State Standards ENTERPRISE

The Common Core State Standards for Mathematics include Standards for Mathematical Practice and Mathematical Content. These standards define the mathematics that students should understand and be able to do. The grade-specific placement of the standards is based on state and international comparisons and the expert opinion of mathematicians and mathematics educators.

The Standards for Mathematical Practice focus on problem solving, reasoning and proof, communication, representation, connections, adaptive reasoning, strategic competence, conceptual understanding, procedural fluency, and productive disposition. These standards identify the ways students engage in mathematical content.

The Standards for Mathematical Content focus on counting and cardinality, operations and algebraic thinking, number and operations in base ten and fractions, geometry, measurement and data, expressions and equations, the number system, functions, ratios and proportional relationships, statistics and probability, algebra, modeling, and number and quantity. The Common Core State Standards Initiative recognizes that "No set of grade-specific standards can fully reflect the great variety in abilities, needs, learning rates, and achievement levels of students in any given classroom. However, the Standards do provide clear signposts along the way to the goal of college and career readiness for all students." (Common Core State Standards for Mathematics 2010, www.corestandards.org/the-standards/mathematics).

STAR Math Enterprise is a K–Grade 12 assessment that focuses on measuring student performance with skills in the following domains: Numbers and Operations; Algebra; Geometry and Measurement; and Data Analysis, Statistics, and Probability. Measures in these areas provide valuable information regarding the acquisition of mathematic ability along the continuum of mathematics expectations.

Resources consulted to determine the set of skills most appropriate for assessing the mathematics development of US students include:

- *Common Core State Standards for Mathematics*
- National Mathematics Advisory Panel *Foundations for Success: The final report of the National Mathematics Advisory Panel*
- National Council of Teachers of Mathematics (NCTM) *Curriculum Focal Points for Prekindergarten Through Grade 8 Mathematics*
- NCTM *Principles and Standards for School Mathematics*
- state standards
- Singapore primary and secondary mathematics standards
- *National Assessment of Educational Progress* (NAEP)
- *Trends in International Mathematics and Science Study* (TIMSS)

Renaissance Learning's Core Progress Learning Progression for Math and the Common Core State Standards

The Common Core State Standards Initiative recognizes the importance of a learning progression of mathematics skills and anticipates that the common state standards will facilitate research in this area. It states, "What students can learn at any particular grade level depends upon what they have learned before. Ideally then, each standard in this document might have been phrased in the form, 'Students who already know... should next come to learn....'" The standards describe this progression from kindergarten through high school. (Common Core State Standards for Mathematics 2010).

Renaissance Learning's researched-based and empirically supported math Core Progress learning progression for math identifies the continuum of math concepts and skills needed for success in math. The continuum begins with early numeracy and progresses through high school algebra and geometry. The skills assessed in STAR Math Enterprise are a subset of this larger continuum of skills. STAR Math Enterprise assessment results are correlated to Renaissance Learning's Core Progress learning progression for math.

Test Length

Each STAR Math session administers 24 test items tailored to the age, grade placement, and actual performance level of the student.

ENTERPRISE The STAR Math Enterprise test administers 34 questions.

Test Administration Time

A STAR Math test typically takes 10–15 minutes to administer. During research and development, about 50 percent of all students finished in less than 12 minutes; 75 percent of all students completed the test in 15 minutes or less.

ENTERPRISE The STAR Math Enterprise test takes slightly longer to administer: 20 minutes is typical.

Adaptive Branching

STAR Math selects items one at a time, based on a continually updated estimate of the student's ability level. Initially, this estimate is based on the student's age and grade placement. Subsequently, it is based on the student's actual performance on previous tests and during the current one. Using Adaptive Branching, the software chooses test items on the basis of content and difficulty, with the objective of matching item difficulty to the student's ability, and producing an average of 75 percent correct (67 percent for STAR Math Enterprise). This Adaptive Branching process is based on the branch of psychometrics called item response theory (IRT).

Test Administration Procedures

In order to ensure consistency and comparability of test results to the STAR Math norms, teachers administering a STAR Math test should follow the recommended administration procedures. These same procedures were used by the norming participants. It is also a good idea to make sure that the testing environment is as free from distractions for the student as possible.

During STAR Math norming, the program was modified so that teachers could not deactivate the proctoring (test-monitoring) options. This was necessary to ensure that the norming data gathered were as reliable as possible. During norming, test monitors had responsibility for test security and were required to provide access to the test for each student. In the final version of the software, teachers can turn off the requirement for test monitoring using the Testing Password preference, but it is not recommended that they do so.

Also during STAR Math norming, all of the participants received the same set of test instructions contained in the STAR Math Pretest Instructions. These instructions describe the standard test orientation procedures that teachers should follow to prepare their students for the STAR Math test. These instructions are intended for use with students of all ages and have been successfully field-tested with students ranging from grade 1 to grade 12. It is important to use these same instructions with all students prior to STAR Math testing. While the Pretest Instructions should be used prior to each student's first STAR Math test, it is not necessary to administer them prior to a student's second or subsequent tests.

Content and Item Development

Content of the STAR Math test evolved through three stages of development. The first stage involved specifying the curriculum content to be reflected in the test. Because rules for writing the items influenced the exact ways in which this content finally appeared in the test, these rules may be considered part of this first stage of development. The following section describes these rules. In the second stage, items were empirically tested in a calibration research program, and items most suited to the test model were retained. The third stage occurs dynamically as each student completes a STAR Math test. The content of each STAR Math test depends on the selection of items for that individual student according to the computer-adaptive testing mode.

Content Specification: STAR Math

STAR Math test content was intended to reflect the objectives commonly taught in the mathematics curriculum of contemporary schools (primarily in the United States). Four major sources helped to define this curriculum content. First, an extensive review of content covered by leading mathematics textbook series was conducted. Second, state curriculum guides or lists of objectives were reviewed. Third, the *Principles and Standards for School Mathematics* of the National Council of Teachers of Mathematics (NCTM) was employed. Finally, content specifications from the National Assessment of Educational Progress (NAEP) and the Trends in International Mathematics and Science Study (TIMSS) were consulted. There is reasonable, although not universal, agreement among these sources about the content of mathematics curricula.

The final STAR Math content specifications were intended to cover the objectives most frequently found in these four sources. In the end, the STAR Math content was organized into eight strands. Two hundred fourteen objectives were then created within these eight strands. Appendix A (page 135) lists the specific objectives in each strand.

Numeration Concepts

The Numeration Concepts strand encompasses 43 objectives, making it the strand with the largest number of objectives. This strand concentrates on conceptual development of the decimal number system. At the lowest levels, it covers cardinal and ordinal numbers through ten (the ones). The strand then proceeds to treatment of the decades (tens), hundreds, thousands, and then larger numbers such as hundred thousands and millions, all in the whole-number realm. At each of these levels of the number system, specific objectives relate to place value identification, number-numeral correspondence, and expanded notation. Following treatment of the whole numbers, the Numeration Concepts strand moves to fractions and decimals. Coverage includes representation of fractions and decimals on the number line, conversions between fractions with different denominators and between fractions and decimals, number-numeral correspondence for decimals, and rounding decimals.

At the highest level, the Numeration Concepts strand encompasses a variety of objectives that could be labeled pre-algebra or simply “advanced concepts.” Included in this category are specific objectives on roots and powers, primes and composites, signed integers, and scientific notation. Because items in the Numeration Concepts strand emphasize understanding basic concepts, they are deliberately written to minimize computational burden.

Computation Processes

The Computation Processes strand includes 39 specific objectives, the second largest number among the STAR Math strands. This strand covers the four basic operations (addition, subtraction, multiplication, and division) with whole numbers, fractions, decimals, and percents. Ratios and proportions are also included in this strand. Coverage of computational skill begins with the basic facts of addition and subtraction, starting with the fact families having sums to 10, then with sums to 18. The strand progresses to addition and subtraction of two-digit and three-digit numbers without regrouping, then with regrouping. At about the same level, basic facts of multiplication and division are introduced. Then, the four operations are applied to more difficult regrouping problems with whole numbers. Fractions are first introduced by way of addition and subtraction of fractions with like denominators. These are relatively easy for students in the middle grades. However, the strand next includes operations with fractions with unlike denominators, mixed numbers, and decimal problems requiring place change, all of which are relatively difficult for students. The Computation Processes strand concludes with a series of objectives requiring operations with percents, ratios, and proportions.

Although the Computation Processes strand can be subdivided into nearly an infinite number of objectives, the STAR Math item bank provides a representative sampling of computational problems that cover the major types of problems students are likely to encounter. Indeed, the item bank does not purport to cover every conceivable computational nuance. In addition, among the more difficult problems involving computation with whole numbers, there are number combinations for which one would ordinarily use a calculator. However, it is expected that students will know how to perform these operations by hand, and hence, a number of such items are included in the STAR Math item bank.

The Numeration Concepts and Computation Processes strands are considered by many to be the heart of the basic mathematics curriculum. Students must know the four operations with whole numbers, fractions, decimals, and percents. Students must know numeration concepts to have an understanding of how the operations work, particularly for regrouping, changing denominators in fractions, and changing places with decimals and percents. As noted above, these two strands constitute the first two thirds of the STAR Math test. Mathematical development within these two strands also serves as the principal basis for instructional recommendations provided in the STAR Math Diagnostic Report.

The remaining strands comprise the latter third of the STAR Math test. This part might be labeled “applications” since many—although not all—of the objectives in this part can be considered practical applications of mathematical content and procedures. It is important to note that research conducted at the item calibration stage of STAR Math development demonstrated that the items in the various strands were strongly unidimensional, thus justifying the use of a single score for purposes of reporting.

Estimation

The Estimation strand is also designed to parallel the Computation Processes strand in terms of the types of operations required. Again, many but not all computation objectives are reflected in this strand. Obviously, in the Estimation strand, students are not required to compute a final answer. With number combinations similar to those represented in the Computation Processes strand, students are asked to estimate an answer. To discourage students from actually computing answers, response options are generally given in round numbers. The range of numerical values used in the options is generally set so that a reasonable estimate is adequate.

Geometry

The Geometry strand in STAR Math begins with simple recognition of plane shapes and their properties. The majority of objectives in the Geometry strand concentrate on the treatment of perimeters and areas, usually covered in the middle grades, and recognition and use of parallels, intersections, and perpendiculars, covered in the middle and junior-high grades. At the more difficult levels, this strand includes application of principles about triangles and the Pythagorean theorem. Other than these latter topics, this strand does not cover the content of the typical college preparatory course in geometry.

Measurement

Although many curricular sources combine geometry and measurement in a single strand, the STAR Math test represents them separately. At the lowest level, the Measurement strand includes objectives on money, temperature, and time (clocks, days of the week, and months of the year). The strand provides coverage of both metric and customary (English) units. Metric objectives include use of the metric prefixes (milli-, centi-, etc.) and the conversion of metric and customary units. The Measurement strand also includes an objective on measurement of angles, one of the best examples of the overlap between the geometry and measurement areas.

Data Analysis and Statistics

This strand begins with simple, straightforward extraction of information from tables, bar charts, and circle graphs. In these early objectives, information needed to answer the question is given directly in the table, chart, or graph. At the next higher level of complexity, students must combine or compare two or more pieces of information in the table, chart, or graph in order to answer the question. This strand also includes several objectives related to probability and statistics. Curricular placement of probability and statistics objectives varies considerably from one source to another. In contrast, using tables, charts, and graphs is commonly encountered across a wide range of grades in nearly all mathematics curricular materials.

Word Problems

The Word Problems strand includes simple, situational applications of computations. In fact, the Word Problems strand is deliberately structured to parallel the Computation Processes strand in terms of the types of operations required.

Most computation objectives are paralleled in the Word Problems strand. For all items in the Word Problems strand, students are presented with a practical problem, and to answer the item correctly, they must determine what type of computational process to use and then correctly apply that process. The reading level of the problems is kept at a low level to ensure valid assessment of ability to solve word problems.

Algebra

The final strand in the curricular structure of the STAR Math item bank is Algebra. Although algebra is generally thought of as a college preparatory course, elements of algebra are actually introduced much earlier than the high school level in the contemporary mathematics curriculum. The use of simple number sentences and the translation of word problems into equations (at a very simple level) are introduced even in the primary grades. Such objectives are included at the lowest level of the STAR Math Algebra strand. The objectives progress rapidly in difficulty to those found in the formal algebra course. These more difficult objectives include operating with polynomials, quadratic equations, and graphs of linear and non-linear functions. See “Appendix C: Algebra Readiness Skills” on page 165.

Objective Clusters

The STAR Math Diagnostic Report contains two bar charts that reflect each student’s performance on the Numeration Concepts and Computation Processes strands. By viewing these two charts, teachers can graphically see how each student is progressing in these two important areas. The STAR Math Diagnostic Report highlights these two strands because they form the foundation for the mathematics curriculum, especially in grades 1–8. According to the National Council of Teachers of Mathematics’ *Principles and Standards for School Mathematics* (NCTM, 2000), “understanding numbers and operations, developing number sense, and gaining fluency in arithmetic computation form the core of mathematics education for the elementary grades” (page 32).

The content in the Numeration Concepts and Computation Processes strands is organized in a hierarchical structure, reflecting the fact that students’ mathematical development (and math curriculum) proceeds in a step-like fashion. In other words, their understanding of harder concepts is dependent upon their understanding the more basic concepts. For example, a student must first learn how to add numbers together before she is able to multiply them.

Because of this hierarchical structure and because every objective within these two strands could not be included on the STAR Math Diagnostic Report, for data reduction purposes, common objectives were grouped together, forming “objective clusters.” Based on the recommendations of a mathematics content expert, the 43 Numeration Concepts objectives and the 39 Computation Processes objectives in STAR Math were grouped into 9 Computation and 8 Numeration clusters. The objectives included in each cluster in each strand are shown in Table 1.

Table 1: STAR Math Strands and Objective Clusters

Strand	Objective Cluster	Objective ID	Objective Name
Numeration Concepts	Ones	N00	Ones: Locate numbers on a number line
		NA1	Ones: Placing numerals in order
		NA2	Ones: Using numerals to indicate quantity
		NA3	Ones: Relate numerals and number words
		NA4	Ones: Use ordinal numbers
	Tens	N01	Tens: Place numerals (10–99) in order of value
		N02	Tens: Associate numeral with group of objects
		N03	Tens: Relate numeral and number word
		N04	Tens: Identify one more/one less across decades
		N05	Tens: Understand the concept of zero
	Hundreds	N06	Hundreds: Place numerals in order of value
		N07	Hundreds: Relate numeral and number word
		N08	Hundreds: Identify place value of digits
		N09	Hundreds: Write numerals in expanded form
	Thousands	N11	Thousands: Place numerals in order of value
		N12	Thousands: Relate numeral and number word
		N13	Thousands: Identify place value of digits
		N14	Thousands: Write numerals in expanded form
	Hundred Thousands	N16	Ten thousands, hundred thousands, millions, billions: Place numerals in order of value
		N17	Ten thousands, hundred thousands, millions, billions: Relate numeral and number word
		N18	Ten thousands, hundred thousands, millions, billions: Identify place value of digits
		N19	Ten thousands, hundred thousands, millions, billions: Write numerals in expanded form

Table 1: STAR Math Strands and Objective Clusters (Continued)

Strand	Objective Cluster	Objective ID	Objective Name
Numeration Concepts (continued)	Fractions & Decimals	N21	Fractions and decimals: Convert fraction to equivalent fraction
		N22	Fractions and decimals: Convert fraction to decimal
		N23	Fractions and decimals: Convert decimal to fraction
		N24	Fractions and decimals: Read word names for decimals to thousandths
		N25	Fractions and decimals: Identify place value of digits in decimals
		N26	Fractions and decimals: Identify position of decimals on number line
		N27	Fractions and decimals: Identify position of fractions on number line
		N28	Fractions and decimals: Convert improper fraction to mixed number
		N29	Fractions and decimals: Round decimals to tenths, hundredths
		N30	Fractions and decimals: Relate decimals to percents
	Advanced Concepts I	N31	Advanced concepts: Determine square roots of perfect squares
		N34	Advanced concepts: Recognize meaning of exponents (2–10)
		N39	Advanced concepts: Can determine greatest common factor
		N41	Advanced concepts: Recognizes use of negative numbers
	Advanced Concepts II	N32	Advanced concepts: Give approximate square roots of a number
		N33	Advanced concepts: Recognize the meaning of n^{th} root
		N35	Advanced concepts: Recognize meaning of negative exponents
		N36	Advanced concepts: Recognize meaning of fractional exponents
		N37	Advanced concepts: Can use scientific notation
		N38	Advanced concepts: Knows meaning of primes and composites
		N40	Advanced concepts: Can determine least common multiple

Table 1: STAR Math Strands and Objective Clusters (Continued)

Strand	Objective Cluster	Objective ID	Objective Name
Computation Processes	Addition & Subtraction Basic Facts to 10	C01	Addition of basic facts to 10
		C02	Subtraction of basic facts to 10
	Addition & Subtraction Basic Facts to 18, No Regrouping	C03	Addition of basic facts to 18
		C04	Subtraction of basic facts to 18
		C05	Addition of three single-digit addends
		C06	Add beyond basic facts, no regrouping (2d + 1d)
		C07	Subtract beyond basic facts, no regrouping (2d – 1d)
	Addition & Subtraction with Regrouping	C08	Add beyond basic facts with regrouping (2d + 1d, 2d + 2d)
		C09	Subtract beyond basic facts with regrouping (2d – 1d, 2d – 2d)
		C10	Add beyond basic facts with double regrouping (3d + 2d, 3d + 3d)
		C11	Subtract beyond basic facts with double regrouping (3d – 2d, 3d – 3d)
	Multiplication & Division: Basic Facts	C12	Multiplication basic facts
		C13	Division basic facts
		C14	Multiplication beyond basic facts, no regrouping (2d × 1d)
	Advanced Computation with Whole Numbers	C15	Division beyond basic facts, no remainders (2d ÷ 1d)
		C16	Multiplication with regrouping (2d × 1d, 2d × 2d)
		C17	Division with remainders (2d ÷ 1d, 3d ÷ 1d)
		C18	Add whole numbers: any difficulty
		C19	Subtract whole numbers: any difficulty
		C21	Divide whole numbers: any difficulty
	Fractions & Decimals I	C22	Add fractions: like single-digit denominators
		C23	Subtract fractions: like single-digit denominators
		C33	Add decimals, place change (2 + .45)
		C35	Subtract decimals, place change (5 – .4)

Table 1: STAR Math Strands and Objective Clusters (Continued)

Strand	Objective Cluster	Objective ID	Objective Name
Computation Processes (continued)	Fractions & Decimals II	C24	Add fractions: unlike single-digit denominators
		C25	Subtract fractions: unlike single-digit denominators
		C26	Multiply fractions: single-digit denominators
		C27	Divide fractions: single-digit denominators
		C28	Add mixed numbers
		C29	Subtract mixed numbers
		C36	Multiply decimals
		C37	Divide decimals
	Percents, Ratios, & Proportions	C38	Percent A (10 is what % of 40?)
		C39	Percent B (20% of 50 is what?)
		C40	Percent C (30 is 50% of what?)
		C41	Proportions
		C42	Ratios
	Multiplication & Division of Mixed Numbers	C30	Multiply mixed numbers
		C31	Divide mixed numbers

On the STAR Math Diagnostic Report, the shaded region of each bar chart reflects the amount of material within each strand that the student has most likely mastered. These estimates are based on the STAR Math norming data, and mastery is defined as 70 percent proficient. Therefore, if a student's ability estimate suggests that she could answer 70 percent or more correct on a specific objective cluster, such as Hundreds, she will have "mastered" that objective cluster and that box will be shaded on her Diagnostic Report. Because the content in the strands included in the objective clusters is hierarchical, students most likely master the objective clusters in sequential order. The solid black line on the bar chart points to the objective cluster that the student is currently developing or the lowest objective that she has not mastered.

Content Specification: STAR Math Enterprise ENTERPRISE

Since STAR Math was introduced in 1998, it has undergone a process of continuous research and improvement. STAR Math Enterprise is an expanded test with new content and several technical innovations. The STAR Math Enterprise item bank was expanded from more than 1,900 test items to more than 4,400 test items. The STAR Math Enterprise Test content was expanded from 210 skills to more than 550 skills to significantly enhance the test's ability to measure math skills in Common Core and other state standards.

For information regarding the development of STAR Math items, see "Item Development Guidelines: STAR Math" on page 23. Before inclusion in the STAR Math Enterprise item bank, all STAR Math items were reviewed to ensure they met the content specifications for STAR Math Enterprise item development. Items that did not meet the specifications were revised and recalibrated. All new item development adheres to the content specifications. All grade 1 through grade 8 items were calibrated using the dynamic calibration method (see page 40). High-school level items used dynamic calibration and fixed-form calibration methods.

The first stage of the expanded STAR Math Enterprise development was identifying the set of skills to be assessed. Multiple resources were consulted to determine the set of skills most appropriate for assessing the mathematics development of K–12 US students, typical mathematics curricula, and current mathematics standards. The resources include:

- *Common Core State Standards for Mathematics*
- National Mathematics Advisory Panel, *Foundations for Success: The final report of the National Mathematics Advisory Panel*
- National Council of Teachers of Mathematics (NCTM), *Curriculum Focal Points for Prekindergarten Through Grade 8 Mathematics*
- NCTM, *Principles and Standards for School Mathematics*
- United States state standards with high quality ratings
- Singapore primary and secondary mathematics standards
- *National Assessment of Educational Progress* (NAEP)
- *Trends in International Mathematics and Science Study* (TIMSS)

The development of the skills list included iterative reviews by mathematicians, mathematics educators, assessment experts, and psychometricians specializing in educational assessment. See "Appendix B: Objectives and STAR Math Enterprise Items" on page 143 for the STAR Math Enterprise Skills List.

The skills list is organized into four domains: Numbers and Operations; Algebra; Geometry and Measurement; and Data Analysis, Statistics, and Probability. To ensure appropriate distribution of items, the assessment blueprint uses six content domains by treating Numbers and Operations and Geometry and Measurement as separate domains.

The second stage included item development and calibration. Assessment 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 writers and editors have content-area expertise and relevant classroom experience and use those qualifications in determining grade-level appropriateness for subject matter and 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. 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, within content strands, to determine whether the item meets all criteria for use in an operational assessment.

STAR Math Enterprise and the Reorganization of Objective Clusters ENTERPRISE

STAR Math Enterprise assesses 550 skills in four standards-based math domains, as outlined in Table 2:

Table 2: Comparison of Domains and Skill Sets: STAR Math vs. STAR Math Enterprise

	STAR Math	STAR Math Enterprise
Skills assessed in:	Eight strands— 1. Numeration 2. Computation 3. Word Problems 4. Geometry 5. Measurement 6. Algebra 7. Estimation 8. Data Analysis and Statistics	Four standards-based domains— 1. Numbers and Operations 2. Algebra 3. Geometry and Measurements 4. Data Analysis, Statistics, and Probability
Skill sets	17	54
Number of skills	210	550

Many of the strands are still represented in the new domains; they are just grouped differently. The reorganization is modeled after the Common Core State Standards.

Within each domain, skills are organized into sets of closely related skills. The resulting hierarchical structure is domain, skill set, and skill. There are four math domains, 54 skill sets, and more than 550 skills. See "Appendix B: Objectives and STAR Math Enterprise Items" on page 143. for a complete list of the STAR Math Enterprise domains, skill sets, and skills.

Item Development Guidelines: STAR Math

When preparing specific items to test student knowledge of the content selected for STAR Math, several item-writing rules were employed. These rules helped to shape the final appearance of the content and hence became part of the content specifications:

- The first and perhaps most important rule was to have the item content, wording, and format reflect the typical appearance of the content in curricular materials. In some testing applications, one might want the item to look different from how the content typically appears in curricular materials. However, the goal for the STAR Math test was to have the items reflect how the content appears in curricular materials that students are likely to have used.

- Second, every effort was made to keep item content simple and to keep the required reading levels low. Although there may be some situations in which one would want to make test items appear complex or use higher levels of reading difficulty, for the STAR Math test, the intent was to simplify when possible.
- Third, efforts were made both in the item-writing and in the item-editing phases to minimize cultural loading, gender stereotyping, and ethnic bias in the items.
- Fourth, the items had to be written in such a way as to be presented in the computer-adaptive format. More specifically, items had to be presentable on the types of computer screens commonly found in schools. This rule had one major implication that influenced item presentation: artwork was limited to fairly simple line drawings, and colors were kept to a minimum.
- Finally, items were all to be presented in a multiple-choice format. Answer choices were to be laid out in either a 4 × 1 matrix, a 2 × 2 matrix, or a 1 × 4 matrix.

In all cases, the distracters chosen were representative of the most common errors for the particular question stem. A "not given" response option was included only for the Computation Processes strand. This option was included to minimize estimation as a response strategy and to encourage the student to actually work the problem to completion.

Item Development Guidelines: STAR Math Enterprise ENTERPRISE

STAR Math Enterprise assesses more than 550 grade-specific skills. 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 Enterprise 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: use of fiction and nonfiction, subject and topic areas, geographic region, gender, ethnicity, occupation, age, and disability.

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 18-point Arial. Graphics are included in an item only when necessary to solve the problem.

Item stems meet the following criteria with limited exceptions. 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 not used as distractors.

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.

Core Progress Learning Progression for Math

STAR Math Enterprise and Core Progress Learning Progression for Math

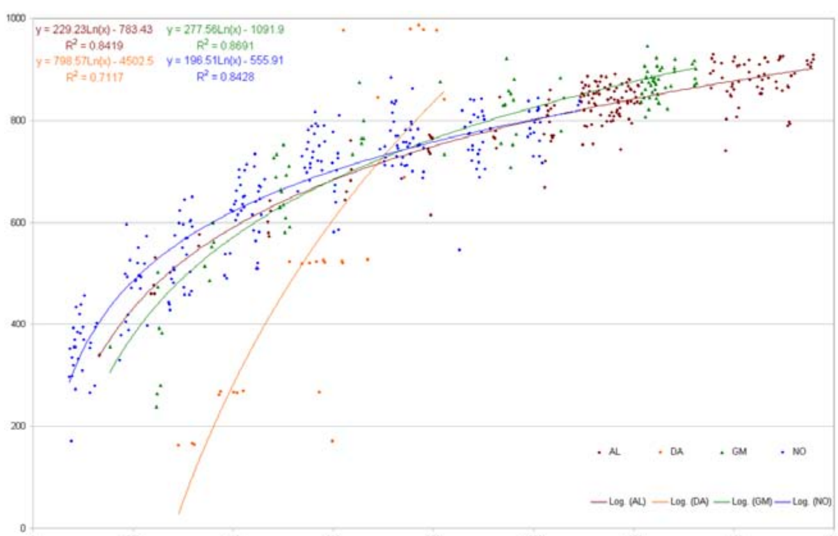
STAR Math Enterprise bridges assessment and instruction through a research-based learning progression for math to help teachers make effective instructional decisions and to adjust instruction to meet the needs of students at different achievement levels.

The learning progression for math identifies the continuum—or instructional sequence—of math concepts and skills spanning from early numeracy through high-school level algebra and geometry. The Core Progress learning progression for math was developed in consultation with leading experts in mathematics and supported by calibration data and psychometric analysis.

The learning progression for math consists of four domains: Numbers and Operations; Algebra; Geometry and Measurement; and Data Analysis, Statistics, and Probability. Within each domain, closely related skills are organized into 23 skill areas. Over 1,300 skills are represented in the skill areas.

To map Core Progress and STAR Math, developers created STAR Math items to assess the skills in the Core Progress learning progression for math. These items were then calibrated to the STAR Math scale, and the skill difficulty was determined from the items. Examination of the item calibration results found that the rank order of the difficulty of the STAR Math Enterprise items correlates closely to the rank order of the skills in the learning progression for math. Figure 1 illustrates the relationship between the instructional order of skills according to the learning progression for math (represented by the trend lines) and the empirical difficulty levels of the skills determined through calibration (represented by the data points).

Figure 1: Skill Difficulty by Domain



This validation process is ongoing. Its purpose is to compare the research-based order of skills against the empirical results of calibration to ensure that the progression in the learning progression for math is an accurate representation of the order in which students learn math skills and concepts. To that end, response data collected from STAR Math Enterprise is continuously used to validate and refine the learning progression for math.

There are two ways to access Core Progress learning progression for math. First, STAR Math Enterprise generates Instructional Planning Reports that use the science of the learning progression for math to identify the range of skills students are ready to learn next. The Instructional Planning reports, coupled with the STAR instructional grouping tool, serve as a starting point for instruction by allowing teachers to differentiate learning and practice opportunities for their students based on the student's scaled score. Second, educators can navigate the learning progression for math through an interactive web portal that includes a search function. Search results provide information about the skills, such as terminology and related skill concepts, as well as prerequisites for learning the skills. The search results also give the teacher access to a variety of instructional resources, including Worked Examples and Sample Questions.

Core Progress learning progression for math is also mapped to the Accelerated Math Second Edition Libraries, enabling teachers to provide appropriate practice activities for their students, including students in need of intervention.

Item and Scale Calibration

Background

The introduction of STAR Math Enterprise marks the third major evolution in the calibration of STAR Math items. For the original version of STAR Math, circa 1997, data for item calibration were collected using printed test booklets and answer sheets, in which the items were formatted to closely match the appearance those items would later take when displayed on computer screens. For STAR Math version 2, data collection was done entirely by computer, using a special-purpose application program that administered fixed test forms, but did so on screen, with the same display format and user interface later used in the adaptive version of STAR Math 2. For STAR Math Enterprise, new test items to be calibrated were embedded as unscored items in STAR Math itself, and the data for calibration were collected by the STAR Math software. Renaissance Learning calls this data collection process *dynamic calibration.*

In the original development of STAR Math (in 1997), approximately 2,450 items were prepared according to the defined STAR Math content specifications. These items were subjected to empirical tryout in a national sample of students in grades 3–12. Following both traditional and item response theory (IRT) analyses of the resulting item response data, 1,434 of the items were chosen for use in the original STAR Math item bank.

In the development of STAR Math 2 in 2002, about 1,100 new items were written. The new items extended the content of the STAR Math item bank to include grades 1–12 and expanded the algebra coverage by adding a number of new algebra objectives. Where needed, items measuring other objectives were written to supplement existing items. (Later versions of the program used this same item bank.)

All of the new items had to be calibrated on the same difficulty scale as the original STAR Math item bank. Because a number of changes in item display features were introduced with STAR Math 2, Renaissance Learning decided to recalibrate the original STAR Math adaptive item bank simultaneously with the new items written specifically for STAR Math 2. During that Calibration Study, 2,471 items, including both the existing and the new items, were administered to a national sample of more than 44,000 students in grades 1–12 in the spring of 2001.

For the development of STAR Math Enterprise, several thousand new items spanning content appropriate for grades 1–10 were developed. Data for calibrating them were collected using the dynamic calibration feature of the Renaissance Place versions of STAR Math. Small numbers of these items were randomly selected for each student, and embedded at appropriate random points in most STAR Math tests administered using Renaissance Place Real Time, beginning in the 2008–2009 school year, and continuing to the present. Each student taking STAR Math on Renaissance Place Real Time was administered a small number of these new, uncalibrated items.

Calibration Samples

The current approaches taken to obtaining examinee samples for STAR Math item calibration are quite different from the approaches taken in the development of item banks for the original STAR Math and STAR Math 2. This section begins with a discussion of the current, dynamic calibration approach. It is followed by a description of the approach taken in the earlier STAR Math 2 calibration.

The Dynamic Calibration Approach

Item calibration entails estimating the scaled difficulty of test items by administering them to examinees whose ability is known or estimated, then fitting 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. With the introduction of dynamic calibration in STAR Math, items to be calibrated are embedded as unscored items in STAR Math tests; to ensure a broad diversity of examinee ability, uncalibrated items are selected randomly and administered to students at the target grade level of each item, as well as one grade level above the target, and in some cases one grade level below.

Although we were not seeking a nationally representative examinee sample, it is useful to evaluate the diversity of the samples who contributed to the calibration data. The tables immediately below describe the overall sample of students who contributed item response data to the calibration of 2,473 STAR Math and STAR Math Enterprise test items over an 18-month period from February 2010 to July 2011. Over 1.5 million students from 7,340 schools in 49 states, in addition to Canada and the US Virgin Islands contributed to the overall response data set. Many of those students took two or more STAR Math tests during that interval; the total number of tests taken was over 3 million. The number of responses per item ranged from 520 to 58,805, with an median of 2,561.

Of the students participating, 1,446,760 were in US schools; selected demographic data on the U.S. students are in the following tables. Table 3 displays the recorded demographic characteristics of those examinees. Table 4 displays the distribution of the examinees by region of the US; examinees from Canada and outside North America also participated, but their numbers were quite small and are not reported here. Table 6 displays the distribution by gender. Entering the data for each of these analyses was optional; each table tallies only those cases for which the relevant data elements were recorded.

Table 3: Sample Ethnicity, STAR Math Calibration Study—February 2010–July 2011 (N = 1,446,760 US Students)

Ethnicity Description	Observations	Observed Percentage	Population Percentage
American Indian or Alaskan Native	16,058	2.99	1.1
Asian or Pacific Islander	16,332	3.04	3.9
Black	156,416	29.13	16.8
Hispanic	105,433	19.64	14.7
Other Race or Ethnicity	1,577	0.29	–
White	241,103	44.90	63.5
Total Observations	536,919		

Table 4: Sample by US Region, STAR Math Calibration Study—February 2010–July 2011 (N = 1,446,760 US Students)

Region	Observations	Observed Percentage	Population Percentage
Midwest	169,311	26.13	23.50
Northeast	39,810	6.14	20.4
Southeast	231,819	35.78	24.30
West	207,042	31.95	31.80
Total	647,982		

Table 5: Sample by Gender, STAR Math Calibration Study—February 2010–July 2011 (N = 1,446,760 US Students)

Gender	Observations	Observed Percentage	Population Percentage
Female	490,357	48.22	Not available
Male	526,471	51.78	
Total	1,016,828		

STAR Math 2 Calibration

To obtain a sample that was representative of the diversity of mathematics achievement in the US school population, school districts, specific schools, and individual students were selected to participate in the STAR Math 2 Calibration Study. The sampling frame consisted of all US schools, stratified on three key variables: geographic region of the country, school size, and socioeconomic status. The STAR Math calibration sample included students from 261 schools from 45 of the 50 United States. Tables 6 and 7 present the characteristics of the calibration sample.

Table 6: Sample Characteristics, STAR Math 2 Calibration Study—Spring 2001 (N = 44,939 Students)

		Students	
		National %	Sample %
Geographic Region	Northeast	20.4%	7.8%
	Midwest	23.5%	22.1%
	Southeast	24.3%	37.3%
	West	31.8%	32.9%
District Socioeconomic Status	Low	28.4%	30.2%
	Average	29.6%	38.9%
	High	31.8%	23.1%
	Non-Public	10.2%	8.1%
School Type and District Enrollment	Public		
	< 200	15.8%	24.2%
	200–499	19.1%	26.2%
	500–1,999	30.2%	26.4%
	2,000 or More	24.7%	15.1%
	Non-Public	10.2%	8.1%

Table 7: Ethnic Group and Gender Participation, STAR Math 2 Calibration Study—Spring 2001 (N = 44,939 Students)

		Students	
		National %	Sample %
Ethnic Group	Asian	3.9%	2.8%
	Black	16.8%	14.9%
	Hispanic	14.7%	10.3%
	Native American	1.1%	1.6%
	White	63.5%	70.4%
	Response Rate	86.2%	35.7%
Gender	Female	Not available	49.8%
	Male	Not available	50.2%
	Response Rate	0.0%	55.9%

Data Collection

STAR Math Enterprise Items ENTERPRISE

Beginning in September 2008, thousands of new, standards-based test items spanning the grade range from grade 1 through Algebra 1 and Geometry were developed, and calibrated by means of analysis of response data collected using the dynamic calibration feature of the STAR Math Renaissance Place versions. Most students taking STAR Math at sites that use Renaissance Place Real Time since that date have had several unscored items embedded among the scored STAR Math test items. The choice of unscored items was done randomly by item grade level; the positions of the unscored items were randomly located, subject to content constraints. Specifically, Numeration Concepts items were embedded among the first 8 scored items; Computation Processes items were embedded among scored items 9 to 16; and items from all other content strands were embedded among scored items 17 to 24. Each STAR Math test recorded the student's final Rasch ability score, based on the 24 scored items, as well as the responses to the unscored items. Unscored items were calibrated on the STAR Math Rasch score scale by calculating the logistic regression of each item's scored responses (0 or 1) on the students' Rasch ability scores. Tables 3, 4, and 5 summarize demographic data on about 1.5 million students and 2,473 items that were part of this process between February 2010 and July 2011. Similar-sized student and item samples were calibrated during other periods, throughout the 2008, 2009, and 2010 school years.

STAR Math 2 Items

The calibration data were collected by administering test items on-screen, with display characteristics identical to those to be implemented in the STAR Math product. However, the calibration items were administered in forms consisting of fixed sequences of items, as opposed to the adaptive testing format.

Seven levels of test forms were constructed corresponding to varying grade levels. Because growth in mathematics is much more rapid in the lower grades, there was only one grade per level for the first four levels. As grade level increases, there is more variation among both students and school curricula, so a single test level can cover more than one grade level. Grades were assigned to test levels after extensive consultation with mathematics instruction experts, and assignments were consistent both with the STAR Math item development framework and with assignments used in other math achievement tests. To create the levels of test forms, therefore, items were assigned to grade levels such that resulting test forms sampled an appropriate range of objectives from each of the strands that are typically represented at or near the targeted grade levels. Table 8 describes the various test form designations used for the STAR Math Calibration Study.

Table 8: Test Form Levels, Grades, Numbers of Items per Form and Numbers of Test Forms, STAR Math 2 Calibration Study—Spring 2001

Level	Grades	Items per Form	Forms	Items
A	1	36	14	152
B	2	36	22	215
C	3	36	32	310
D	4	36	34	290
E	5–6	46	36	528
F	7–9	46	32	516
G	10–12	46	32	464

Students in grades 1–4 (Levels A, B, C, and D) took 36-item tests consisting of three practice items and 33 actual test items. Expected testing time for these students was 30 minutes. Students in grades 5–12 (Levels E, F, and G) took 46-item tests consisting of three practice items and 43 actual test items. Expected testing time for these students was 40 minutes.

Items within each level were distributed among a number of test forms. Consistent with the previous version of STAR Math, the content of each form was balanced between two broad categories of items: items measuring Numeration Concepts and Computation Processes and items measuring Other Applications. Each form was organized into three sections: A, B, and C. Sections A and C each consisted of approximately 40% of the test length, and contained items from both of the categories.

Section A began with items measuring Numeration Concepts and Computation Processes, followed by items measuring Other Applications. Section C reversed this order, with Other Applications items preceding Numeration Concepts and Computation Processes items.

Section B comprised approximately 20% of the test length, and contained two types of anchor items. “Horizontal anchors” were common to a number of test forms at the same level, and “vertical anchors” were common to forms at adjacent levels. The anchor items were used to facilitate later analyses that placed all item difficulty parameters on a common scale.

With the exception of Levels A and G, approximately half of the vertical anchor items in each form came from the next lower level, and the other half came from the next higher level. Items chosen as vertical anchor items were selected partially based on their difficulty; items expected to be answered correctly by more than 80 percent or fewer than 50 percent of out-of-level students were not used as vertical anchor items. Two versions of each form were used: version A and version B. Each version A form consisted of Sections A, B, and C in that order. Each version B form contained the same items, arranged in reverse order, with Section C followed by Sections B and A. The alternate forms counterbalanced the order of item presentation, as a defense against possible order effects influencing the psychometric properties of the items. In all three test sections, items were chosen so that content was balanced at each level, with the numbers of items measuring each of the content domains roughly proportional to the distribution of items among the domains at each level.

In Levels A–G combined, there were 101 unique sets of test items. Each was arranged in two alternate forms, versions A and B, that differed only in terms of item presentation order. Therefore, there was a total of 202 test forms.

Item Analysis

Both STAR Math Enterprise and STAR Math 2 analyses followed similar courses. Following extensive quality control checks, the item response data were analyzed by level, using both traditional item analysis techniques and item response theory (IRT) methods. For each test item, the following information was 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 option and the alternatives.
- 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.

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 Discrimination

The traditional measure of the discrimination 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 2 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 different from 0.50, and particularly as the p-value approaches 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 Enterprise have used the Urry biserial as the correlation coefficient of choice; item selection/rejection decisions have been based in part on his suggested target of 0.62.

Item Response Function

In addition to traditional item analyses, the STAR Math calibration data were analyzed using item response theory (IRT) methods. Item response theory is widely recognized as the most sophisticated testing approach today.

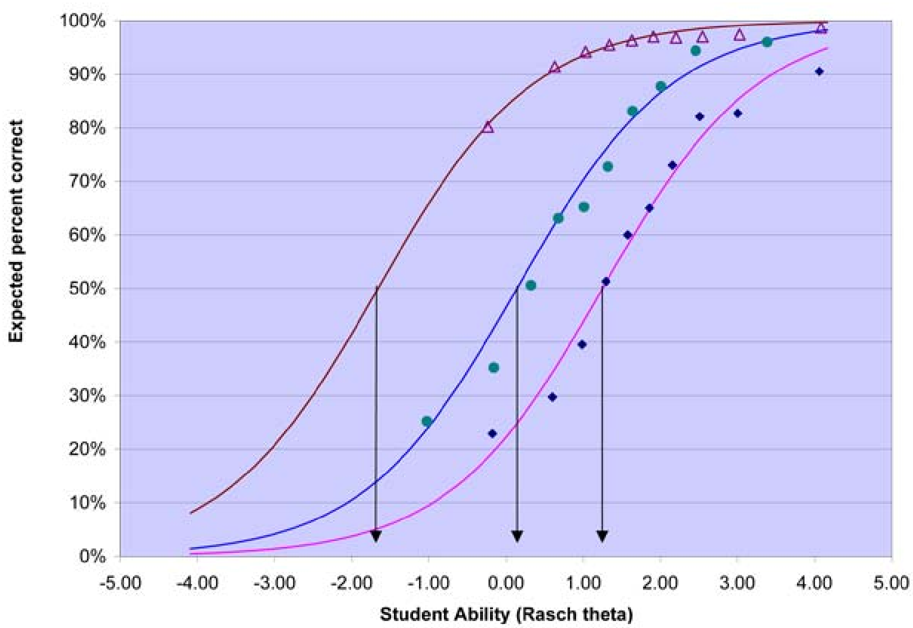
With IRT, the performance of students and the items they answer are placed on the same scale. To accomplish this, every test question is calibrated. Calibration is a research-based method for determining the difficulty of a test question. It is done by administering each question to hundreds and sometimes 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 easiest to hardest question in the item bank. After taking a STAR assessment, a student's score is plotted on this developmental scale. Placing students and items on the same scale is the breakthrough of IRT because it makes it possible to assign scores on the same scale even though students take different tests. IRT also provides a means to identify 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 data both for its simplicity and its ability to accurately model the performance of the STAR Math 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 2 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 2: Three Examples of Item Response Functions



Item response theory expresses both item difficulty and student ability on the same scale. In Figure 2, 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 difficulty parameter for each test item and places all of the item parameters onto a common scale. 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 2 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.

Review of Calibrated Items

Following these analyses, each test item, along with both traditional and IRT analysis information (including IRF and EIRF plots), and information about the test level, form, and item identifier were 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 or STAR Math Enterprise 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, did not meet acceptable standards for inclusion in the STAR Math item bank.

Rules for Item Retention

Items were eliminated if any of the following occurred:

- STAR Math Enterprise: The Urry biserial correlation (item discrimination was less than 0.62.
- STAR Math 2: The item-total correlation (item discrimination) was less than 0.30.
- At least one of an item's distracters had a positive item discrimination.
- The sample size of students attempting the item was less than 300.
- The traditional item difficulty indicated that the item was too difficult or too easy.
- The item did 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 Enterprise item calibration above, 884 items (36%) met all the retention rules above, and were accepted for operational use as part of the STAR Math enterprise 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 Enterprise.

In the case of the STAR Math 2 items, of the initial 2,471 items administered in the STAR Math Calibration Study, approximately 2,000 (81%) were deemed of sufficient quality to be retained for further analyses. About 1,200 of these retained items were items from the original version of STAR Math. Traditional item-level analyses were conducted again on the reduced data set. In these analyses, the dimensionality assumption of combining the first and second parts of the test was re-evaluated to ensure that all items could be placed onto a single scale. In the final IRT calibration, all test forms and levels were equated based on the information provided by the embedded anchor items within each test form so that the resulting IRT item difficulty parameters were placed onto a single scale spanning grades 1–12.

Computer-Adaptive Test Design

An additional level of content specification is determined by the student's performance during testing. In conventional paper-and-pencil standardized tests, items retained from the item tryout or item calibration program are organized by level. Then, each student takes all items within a given test level. Thus, the student is only tested on those mathematical operations and concepts deemed to be appropriate for his or her grade level.

On the other hand, in computer-adaptive tests, such as STAR Math, the items taken by a student are dynamically selected in light of that student's performance during the testing session. Thus, a low-performing student's knowledge of math operations may branch to easier operations to better estimate math achievement level, and 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 responded to 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 STAR Math, the adaptive branching procedure aims to select items such that a student is expected to have a 75 percent chance of answering each item correctly, given the student's estimated ability and the item's known difficulty. In the case of STAR Math Enterprise, the brancher selects items with a 67 percent expectation of a correct response. STAR Math item difficulties were determined by results of the national item Calibration Study.

A STAR Math test consists of a fixed-length, 24-item adaptive test (34 items for STAR Math Enterprise). 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 described in "Conversion Tables" on page 128.

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 this estimated achievement level. Randomization of items with difficulty values near the student's math achievement level allows the program to avoid overexposure of test items.

In the case of STAR Math, the items in the first part of the test (items 1–16) are dynamically selected from an item bank consisting of all the retained items from the Numeration Concepts and Computation Processes strands. Although the second part of the test selects items from a pool that consists of the remaining six content strands, content balancing rules ensure that every strand appropriate to the student's grade level is represented. Table 9 on the next page shows the content balancing design of STAR Math strands by grade.

Table 9: Content-Balancing Design of STAR Math's Strands by Grade–Minimum Distribution of Items by Strands

Strand	Grade 1	2	3	4	5	6	7	8	9	10	11	12	
Computation Processes	8	8	8	8	8	8	8	8	8	8	8	8	First 16 Items (1–16)
Numeration Concepts	8	8	8	8	8	8	8	8	8	8	8	8	
Total	16	16	16	16	16	16	16	16	16	16	16	16	
Strand	**Grade 1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	
Algebra	0	0	0	0	0	0	0	0	2	2	2	2	Last 8 Items (17–24)
Data Analysis and Statistics	1	1	1	1	1	1	1	1	1	1	1	1	
Estimation[a]	–	–	1	1	1	1	1	1	0	0	0	0	
Geometry	2	2	1	1	1	2	2	2	2	2	2	2	
Measurement	2	2	2	2	2	1	1	1	1	1	1	1	
Word Problems	2	2	2	2	2	2	2	2	1	1	1	1	
Total	7	7	7	7	7	7	7	7	7	7	7	7	

a. Students in kindergarten through grade 2 will not receive items from the Estimation strand.

As can be seen in Table 9, all students in all grades receive eight items from Computation Processes and eight items from Numeration Concepts during the first sixteen items of the test. The specific type of question administered within these strands will vary with the student's grade level and estimated ability level. The next seven items are selected according to the student's grade level, according to Table 9. A zero means that no minimum criterion exists, but students may receive items from that strand if it would be consistent with the software's estimated ability level. The final and 24th item of a STAR Math test will be selected from any available strands in Other Applications that are consistent with the student's estimated ability level.

Items that have been administered to the same student within the past 75 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.

STAR Math Scoring

Following the administration of each STAR Math and STAR Math Enterprise 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. Accompanying each Scaled Score is an associated measure of the degree of uncertainty, called the standard error of measurement (SEM). Unlike conventional paper-and-pencil tests, the SEM values for Scaled Scores will be unique for each student dependent upon the particular items in the student's individual test and the student's performance on those items. Because the test is computer-adaptive, however, the SEM values are relatively consistent by the end of the test.

Scaled Scores are expressed on a common scale that spans all grade levels covered by the STAR Math test. Because the software expresses Scaled Scores on a common scale, Scaled Scores are directly comparable with each other, regardless of grade level. Other scores, such as Percentile Ranks and Grade Equivalents, are derived from the Scaled Scores obtained during the STAR Math norming study described in "Conversion Tables" on page 128.

Dynamic Calibration

This feature allows response data on new test items to be collected during the STAR testing sessions for the purpose of field testing and calibrating those items. When dynamic calibration is active, it works by embedding one or more new items at random points during a STAR test. These items do not count toward the student's STAR test score, but item responses are stored for later psychometric analysis. Students may take as many as five additional items per test; in some cases, no additional items will be administered. On average, this will only increase testing time by one to two minutes. The new, non-calibrated items will not count toward students' final scores, but will be analyzed in conjunction with the responses of hundreds of other students.

Student identification does not enter into the analyses; they are statistical analyses only. The response data collected on new items allows for continual evaluation of new item content and will contribute to continuous improvement in STAR tests' assessment of student performance.

Reliability and Measurement Precision

Reliability is a measure of the degree to which test scores are consistent across repeated administrations of the same or similar tests to the same group or population. To the extent that a test is reliable, its scores are free from errors of measurement. In educational assessment, however, some degree of measurement error is inevitable. One reason for this is that a student's performance may vary from one occasion to another. Another reason is that variation in the content of the test from one occasion to another may cause scores to vary.

In a computer-adaptive test such as STAR Math, content varies from one administration to another, and it also varies according to the level of each student's performance. Another feature of computer-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 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, the CSEM is an estimate of the unreliability of each individual test score. While a reliability coefficient is a single value that applies to the overall test, the magnitude of the CSEM may vary substantially from one person's test score to another.

This chapter presents three different types of reliability coefficients: generic reliability, split-half reliability, and alternate forms reliability. This is followed by statistics on the conditional standard error of measurement of STAR Math test scores.

The reliability and measurement error presentation is divided into two sections below: First is a section describing the reliability coefficients and conditional errors of measurement for the original 24-item STAR Math test. Second, another brief section presents reliability and measurement error data for the new, 34-item STAR Math Enterprise test.

24-Item STAR Math Test

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:

$$\text{reliability} = 1 - \frac{\sigma^2_{error}}{\sigma^2_{total}}$$

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, 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.

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

where the summation is over the squared values of the reported CSEM for students i = 1 to n. In each STAR Math 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.

Using this technique with the STAR Math norming data resulted in the generic reliability estimates shown in the rightmost column of Table 10 on page 44. Because this method 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 computer-adaptive test.

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) cannot be calculated for adaptive tests. However, an estimate of internal consistency reliability can be calculated using the split-half method. This is discussed in the next section.

Split-Half Reliability

In classical test theory, before the advent of digital computers automated the calculation of internal consistency reliability measures such as Cronbach's alpha, approximations such as the split-half method were sometimes used. 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.

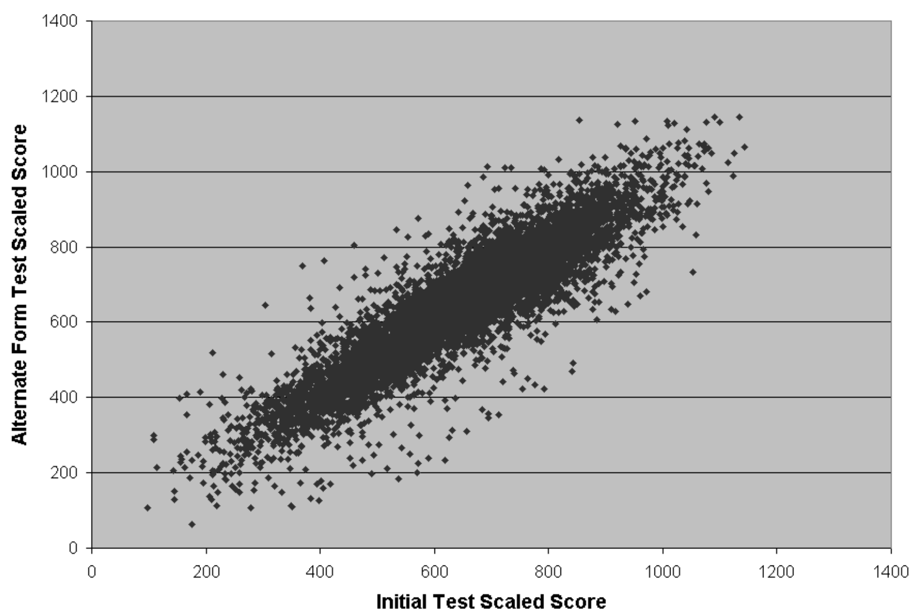
In internal simulation studies, the split-half method provided accurate estimates of the internal consistency reliability of adaptive tests, and so it has been used to provide estimates of STAR Math reliability. These split-half reliability coefficients are independent of the generic reliability approach discussed above and more firmly grounded in the item response data. The fifth column of Table 10 on page 44 contains split-half reliability estimates for STAR Math, calculated from the norming study data.

Alternate Form Reliability

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

This approach was used for STAR Math, as part of the norming study, and the results are presented in the third column of Table 10 on page 44. Participating schools were asked to administer two norming tests, each on a different day, to about one fourth of the overall sample. Figure 3 is a scatterplot of their scores. This resulted in an alternate forms reliability subsample of more than 7,000 students who took different forms of the 24-item STAR Math norming test. The interval between the first and second tests averaged four days. The interval varied widely, however. For example, in some cases both tests were given on the same day; in other cases, the interval ranged from one to as many as 40 days.

Figure 3: Scatterplot of Test Scores from the STAR Math Norming Alternate Forms Reliability Study



Errors of measurement due to both content sampling and temporal changes in individuals' performance can affect alternate forms reliability coefficients, usually making them appreciably lower than internal consistency reliability coefficients. In addition, any growth in the trait that takes place in the interval between tests can also lower the correlation. The actual reliability of STAR Math is probably higher than the alternate forms estimates presented in Table 10 on page 44. Table 10 lists the detailed results of the generic, split-half, and alternate forms reliability analyses of STAR Math Scaled Scores (from the norming study), both overall and by grade.

The split-half and generic reliability estimates, which are based on the entire STAR Math norms sample of 29,228 students,[3] are very similar to one another, with the split-half values generally slightly lower. In the overall sample, these reliability estimates were approximately 0.94. By grade, they range from 0.78 to 0.88, with a median of 0.85.

The alternate forms reliability estimates are based on the 7,517 students who participated in the reliability study, about one fourth of the norms sample. In the overall sample, the alternate forms reliability estimates were approximately 0.91. By grade, the values ranged from approximately 0.72 to 0.80, with a median value of 0.74.

Table 10: Reliability Estimates by Grade from the Norming Study—STAR Math Scaled Scores

Grade	N	Alternate Forms Reliability	N	Split-Half Reliability	Generic Reliability
1	745	0.731	3,076	0.824	0.834
2	866	0.753	3,193	0.777	0.790
3	853	0.741	2,972	0.781	0.798
4	840	0.733	2,981	0.790	0.813
5	813	0.789	3,266	0.803	0.826
6	729	0.734	2,555	0.836	0.838
7	698	0.721	2,896	0.857	0.864
8	714	0.736	2,598	0.877	0.876
9	381	0.793	1,771	0.856	0.862
10	304	0.799	1,556	0.874	0.877
11	255	0.756	1,419	0.865	0.868
12	191	0.722	945	0.882	0.872
Overall	7,389	0.908	29,228	0.944	0.947

Standard Error of Measurement

When interpreting any educational test scores, the test user must bear in mind that the scores include some degree of error. The size of the test score reliability coefficient provides an indication of the overall magnitude of that error. The standard error of measurement (SEM) arguably provides a measure that is more useful for score interpretation, as the SEM is expressed in the same units used to express the test score. For the STAR Math Scaled Score, a conditional SEM is calculated for each individual, and the value of the SEM is included in the score reports, either explicitly or graphically.

3. There were 29,228 cases in the STAR Math 2.0 norms sample; 43 with outlier scores were not included in the norms calculations, but were included in the reliability calculations.

In the following section, aggregate SEMs are presented. For the Scaled Score, these SEMs represent averages, overall and by grade. Because the conditional SEMs vary systematically by Scaled Score, the individual SEMs in the STAR Math score reports are more useful for score interpretation; the averages presented here are for purposes of test evaluation.

Scaled Score SEMs

The STAR Math software calculates the SEM for each individual. This statistic is called the "conditional SEM" as it is conditional on the value of the Scaled Score. Conditional SEMs vary from one student to another, and the interpretation of individual scores should be based on the student's own CSEM value. However, for purposes of summarizing the measurement precision of STAR Math, average conditional SEM values are in Table 11. As the CSEM estimates may vary with ability level, these SEM estimates will be tallied separately for each grade, as well as overall.

Table 11 contains means and standard deviations of the STAR Math Scaled Score conditional SEMs, overall and by grade, for the STAR Math norms sample. The aggregate mean SEM value was 40, averaged over all grades. Within-grade averages range from 37 at grade 1 to 42 at grade 12.

Table 11: STAR Math Standard Error of Measurement of Scaled Scores

		Conditional SEM	
Grade	N	Mean	S.D.
1	3,076	37	5.1
2	3,193	40	4.6
3	2,972	39	3.8
4	2,981	39	3.9
5	3,266	41	4.5
6	2,555	41	4.9
7	2,896	41	5.1
8	2,598	41	5.5
9	1,771	41	5.6
10	1,556	42	6.4
11	1,419	42	6.0
12	945	42	6.6
Overall	29,228	40	5.2

34-Item STAR Math Enterprise Test ENTERPRISE

Reliability Coefficients

STAR Math Enterprise was designed to be the first standards-based STAR assessment, meaning that its item bank measures skills identified by exhaustive analysis of national and state standards in math, from grade K through Algebra I and Geometry. SME content covers almost three times as many skills as previous editions of STAR Math.

Additionally, STAR Math Enterprise items were selected on the basis of the most stringent criteria for technical quality ever applied by Renaissance Learning.

The increased length of STAR Math Enterprise, combined with its increased breadth of skills coverage and enhanced technical quality is expected to result in greater validity than ever before; this should be reflected in higher correlations between STAR Math and other tests, such as state accountability tests. Another expected result is improved measurement precision; this will show up as increased reliability—both internal consistency reliability and test-retest reliability.

Analysis of the first 14,000 SME tests, administered in April 2011, has provided us with early data on the internal consistency reliability of STAR Math Enterprise. Table 12 displays the estimated reliability of SME tests by grade; compare Table 12 to Table 10 for a comparison of the new test's reliability against that of earlier versions of STAR Math.

Table 12: Reliability Estimates by Grade for STAR Math Enterprise

Grade	Sample Size	Internal Consistency Reliability[a]	
		Split-Half Reliability	Generic Reliability
K	53	0.863	0.900
1	1,425	0.887	0.898
2	1,560	0.908	0.915
3	1,791	0.921	0.927
4	2,223	0.929	0.935
5	2,432	0.938	0.941
6	1,533	0.948	0.951
7	1,213	0.935	0.943
8	876	0.947	0.952
9	439	0.926	0.927
10	148	0.968	0.967
11	112	0.931	0.930
12	211	0.924	0.937

a. Reliability estimated using the split-half method.

As Table 12 shows, STAR Math Enterprise reliability is appreciably higher, grade by grade, than the shorter Classic and Service versions. The Enterprise version takes STAR Math to new heights in technical quality, putting this interim assessment on a virtually equal footing with the highest quality summative assessments in use today.

Standard Error of Measurement

Table 13 contains two different sets of estimates of STAR Math Enterprise measurement error: conditional standard error of measurement (CSEM) and global standard error of measurement (SEM). Conditional SEM was described earlier in the introduction of this section on Reliability and Measurement Precision; the estimates of CSEM in Table 13 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)

Global estimates of SEM can be expected to be more conservative (larger) than CSEM estimates, because the former are calculated from observed data, while the individual CSEM values are theory-based. To the extent that students' item responses do not perfectly fit the IRT model used (here, the Rasch model), CSEM should underestimate measurement error. Consistent with that, Table 13's global values of SEM are equal to or greater than the counterpart CSEM values at every grade. However CSEM and SEM are no more than one Scaled Score point different from one another for grades 1 through 12. Only at grade K do they differ by more than one point. The similarity of the values provides confidence that these estimates of SME measurement error are reasonably accurate.

Comparing the estimates of reliability and measurement error of STAR Math (Tables 10, 11) with those of STAR Math Enterprise (Tables 12, 13) confirms that STAR Math Enterprise is appreciably superior to the shorter STAR Math assessments in terms of reliability and measurement precision.

Table 13: Estimates of STAR Math Enterprise Measurement Precision by Grade: Conditional and Global Standard Error of Measurement

Grade	Sample Size	Conditional Standard Error of Measurement Average CSEM	Standard Deviation	Global Standard Error of Measurement (SEM)
K	53	31	7.1	37
1	1,425	30	2.4	32
2	1,560	30	2.6	31
3	1,791	30	3.1	31
4	2,223	30	2.2	31
5	2,432	30	3.1	31
6	1,533	30	2.6	31
7	1,213	30	2.1	32
8	876	30	3.1	32
9	439	30	2.4	30
10	148	31	3.2	31
11	112	30	1.6	30
12	211	30	3.2	33
Average		30		32

The National Center on Response to Intervention (NCRTI) and Progress Monitoring

NCRTI is a federally-funded project whose mission includes reviewing the technical adequacy of assessments as screening and/or progress-monitoring tools for use in schools adopting multi-tiered systems of support (commonly known as RTI, or response to intervention). STAR Math is one of a very small number of mathematics assessments that was judged by NCRTI as being appropriate for both screening and progress monitoring. As of July 2011, STAR Math had the strongest ratings on NCRTI's technical criteria of all mathematics assessments for screening and progress monitoring.

This section highlights results of analyses reviewed by NCRTI related to its progress monitoring domain. For the progress monitoring domain, NCRTI requests information on:

- reliability of the performance level score
- reliability of the slope
- validity of the performance level score
- predictive validity of the slope of improvement
- disaggregated reliability and validity data

For each of these categories, NCRTI assigns one of four qualitative labels: convincing evidence, partially convincing evidence, unconvincing evidence, or data unavailable/inadequate. Please refer to Table 14 for descriptions of these categories as provided by NCRTI. In addition, Table 15 provides the scores assigned to STAR Math in each of the noted categories. Tables 16–20 provide reliability and validity data used to assign the scores outlined below. Further descriptive information is provided within each table.

Table 14: NCRTI Progress Monitoring Indicator Descriptions

Indicator	Description	STAR Math Score
Reliability of the Performance Level Score	Reliability of the performance level score is the extent to which the score (or average/median of 2–3 scores) is accurate and consistent.	Convincing Evidence
Reliability of the Slope	Reliability of the slope is an indicator of how well individual differences in growth trajectories can be detected using a particular measure.	Convincing Evidence
Validity of the Performance Level Score	Validity of the performance level score is the extent to which the score (or average/median of 2–3 scores) represents the underlying construct.	Convincing Evidence
Predictive Validity of the Slope of Improvement	Validity of the slope of improvement is the extent to which the slope of improvement corresponds to end-level performance on highly valued outcomes.	Convincing Evidence
Disaggregated Reliability and Validity Data	Disaggregated data are scores that are calculated and reported separately for specific sub-groups (e.g., race, economic status, special education status, etc.).	Convincing Evidence

Table 15: Reliability of the Performance Level Score for STAR Math

Type of Reliability	Grade	N (Range)	Coefficient		SEM	Information (Including Normative Data)/Subjects
			Range	Median		
Generic	1–5	2,972–3,266	0.790–0.834	0.813	Mean Range 37–41	Based on STAR Math 2.0 norms sample, IRT reliability was calculated from the conditional error variance of IRT ability estimates.
Split Half	1–5	2,972–3,266	0.777–0.824	0.790	NA	Split-half reliability was calculated with the same sample as generic reliability.
Retest	1–5	745–866	0.731–0.789	0.741	NA	There were no common items across retests; non-overlapping versions of STAR Math were taken.
Generic	6–12	945–2,896	0.838–0.877	0.868	Mean Range 41–42	Based on STAR Math 2.0 norms sample, IRT reliability was calculated from the conditional error variance of IRT ability estimates.
Split Half	6–12	945–2,896	0.836–0.882	0.865	NA	Split-half reliability was calculated with the same sample as generic reliability.
Retest	6–12	191–729	0.721–0.799	0.736	NA	There were no common items across retests; non-overlapping versions of STAR math were taken.

Table 16: Reliability of the Slope for STAR Math

Type of Reliability	Grade	N (Range)	Coefficient	Information (Including Normative Data)/Subjects
Split-Half	3	16,651	0.71	Reliability of slope was computed using STAR Math data from school year 2007/08 based on the method recommended in the NCRTI's Frequently Asked Questions document (dated 10/15/2008) and also as described in VanDerHeyden, A., & Burns, M. (2008).
	4	17,187	0.71	
	5	15,310	0.71	
	6	10,026	0.70	
	7	6,205	0.69	
	8	4,878	0.71	
	3	4,894	0.73	Reliability of slope was computed using STAR Math data from school year 2005-06/2006-07 based on the method recommended in the NCRTI's Frequently Asked Questions document (dated 10/15/2008) and also as described in VanDerHeyden, A., & Burns, M. (2008).
	4	5,254	0.74	
	5	2,164	0.74	
	6	1,474	0.69	
	7	1,191	0.72	
	8	127	0.76	

Table 17: Validity of the Performance Level Score for STAR Math

Type of Validity	Grade	Criterion	N (Range)	Coefficient		Information (Including Normative Data)/Subjects
				Range	Median	
Concurrent	1–12	Various	10,000+	0.63–0.65	0.64	Meta-analysis of the 276 correlations with other tests done during the STAR Math 2.0 pilot study were combined and analyzed using a fixed effects model.
Predictive	1–6		11,800–55,285	0.55–0.73	0.67	STAR Math scores predicting later performance on tests including DSTP, FCAT, MEAP, MCA, MCT, NWEA NALT & MAP, OCCT, SM, TAAS, TAKS, Terra Nova (avg. validity).
Predictive	7–12		885–18,919	0.75–0.80	0.76	STAR Math scores predicting later performance on tests including DSTP, MEAP, OCCT, SM, TAAS, TAKS (avg. validity).
Concurrent	3–8		2,335–4,372	0.62–0.70	0.66	STAR Math correlations with State Accountability Tests including DSTP, FCAT, ISAT, MEAP, MCA, MCT, OCCT, TAAS, TAKS (avg. validity).
Predictive	3–8		1,457–1,955	0.49–0.70	0.62	STAR Math scores predicting performance on State Accountability Tests including DSTP, FCAT, MEAP, MCA, MCT, OCCT, TAAS, TAKS (avg. validity).

Table 18: Predictive Validity of the Slope of Improvement for STAR Math

Type of Validity	Grade	Test	Sample Size		Coefficient		Information (Including Normative Data)/Subjects
			Range	Total	Range	Median	
Predictive	3	State Assessment	5–176	529	0.27–0.95	0.65	School years included 2006–07, 2007–08, 2008–09, and 2009–10. STAR Math slopes were correlated with Mississippi, North Carolina, and Oregon End-of-Grade Tests (MCT2, NC EOG, OAKS). Analyses were performed within decile based on starting STAR Math score.
	4		10–91	457	0.33–0.74	0.59	
	5		36–51	394	0.31–0.56	0.46	
	6		8–21	151	0.17–0.78	0.53	
	7		9–15	43	0.57–0.76	0.67	
	8		6–8	14	0.71–0.79	0.74	

Table 19: Disaggregated Validity of the Performance Level Score for STAR Math

Type of Reliability	Age or Grade	N (Range)	Coefficient		SEM
			Range	Median	
Generic (White)	Grades 1–5	33,011	0.81–0.86	0.83	38
Generic (Black)		14,782	0.83–0.89	0.85	38
Generic (Hispanic)		18,450	0.81–0.89	0.86	38
Generic (White)	Grades 6–12	14,991	0.88–0.93	0.90	38
Generic (Black)		7,024	0.90–0.93	0.91	38
Generic (Hispanic)		9,781	0.90–0.93	0.91	38

Table 20: Disaggregated Reliability of the Slope

Type of Reliability	Age or Grade	N (Range)	Median Coefficient	Information (Including Normative Data)/Subjects
Split-Half (Black)	3	747	0.72	Reliability of slope was computed using STAR Math data from school year 2007/08 based on the method recommended in the NCRTI's Frequently Asked Questions document (dated 10/15/2008) and also as described in VanDerHeyden, A., & Burns, M. (2008).
Split-Half (Hispanic)		892	0.71	
Split-Half (White)		2,314	0.69	
Split-Half (Black)	4	648	0.70	
Split-Half (Hispanic)		951	0.69	
Split-Half (White)		2,192	0.71	
Split-Half (Black)	5	621	0.72	
Split-Half (Hispanic)		948	0.69	
Split-Half (White)		2,258	0.71	
Split-Half (Black)	6	388	0.76	
Split-Half (Hispanic)		671	0.71	
Split-Half (White)		1,664	0.73	
Split-Half (Black)	7	394	0.77	
Split-Half (Hispanic)		601	0.71	
Split-Half (White)		1,227	0.67	
Split-Half (Black)	8	275	0.72	
Split-Half (Hispanic)		413	0.76	
Split-Half (White)		1,009	0.71	

The National Center on Response to Intervention (NCRTI) and Screening

For the screening domain, NCRTI requests information on:

- classification accuracy
- reliability
- validity
- disaggregated reliability, validity, and classification data for diverse populations

For each of these categories, NCRTI assigns one of four qualitative labels: convincing evidence, partially convincing evidence, unconvincing evidence, or data unavailable/inadequate. Please refer to Table 21 for descriptions of these categories as provided by NCRTI. In addition, Table 22 provides the scores assigned to STAR Math in each of the noted categories. Tables 23–24 provide the reliability and validity information used to evaluate STAR Math. Further descriptive information is provided within each table.

Table 21: NCRTI Screening Indicator Descriptions

Indicator	Description	STAR Math Score
Classification Accuracy	Classification accuracy refers to the extent to which a screening tool is able to accurately classify students into “at risk for reading disability” and “not at risk for reading disability” categories (often evidenced by AUC values greater than 0.85).	Partially Convincing Evidence
Reliability	Reliability refers to the consistency with which a tool classifies students from one administration to the next. A tool is considered reliable if it produces the same results when administering the test under different conditions, at different times, or using different forms of the test (often evidence by reliability coefficients greater than 0.80).	Convincing Evidence
Validity	Validity refers to the extent to which a tool accurately measures the underlying construct that it is intended to measure (often evidenced by coefficients greater than 0.70).	Convincing Evidence
Disaggregated Reliability, Validity, and Classification Data for Diverse Populations	Data are disaggregated when they are calculated and reported separately for specific sub-groups.	Convincing Evidence

Aggregated Classification Accuracy Data

Receiver Operating Characteristic (ROC) Curves as defined by NCRTI:

> “Receiver Operating Characteristic (ROC) curves are a useful way to interpret sensitivity and specificity levels and to determine related cut scores. ROC curves are a generalization of the set of potential combinations of sensitivity and specificity possible for predictors.” (Pepe, Janes, Longton, Leisenring, & Newcomb, 2004)

> “ROC curve analyses not only provide information about cut scores, but also provide a natural common scale for comparing different predictors that are measured in different units, whereas the odds ratio in logistic regression analysis must be interpreted according to a unit increase in the value of the predictor, which can make comparison between predictors difficult.” (Pepe, et al., 2004)

> “An overall indication of the diagnostic accuracy of a ROC curve is the area under the curve (AUC). AUC values closer to 1 indicate the screening measure reliably distinguishes among students with satisfactory and unsatisfactory reading performance, whereas values at .50 indicate the predictor is no better than chance.” (Zhou, X. H., Obuchowski, N. A., & Obushcowski, D. M., 2002)

Brief Description of the Current Sample and Procedure

STAR Math classification analyses were performed using state assessment data from Arkansas, Delaware, Illinois, Michigan, Mississippi, Oklahoma, Kansas, and North Carolina. Collectively these states cover most regions of the country (Central, Southwest, Northeast, Midwest, and Southeast). The classification accuracy and cross validation study samples were drawn from an initial pool of 29,594 matched student records covering grades 2–8.

The sample used for this analysis was 35% female and 35.1% male, with 29.9% not responding. 24% of students were White, 12.3% were Black, non-Hispanic, and 2.1% were Hispanic. Lastly, 0.3% were Asian or Pacific Islander and 0.4% were American Indian or Alaskan Native. Ethnicity data were not provided for 60.9% of the sample.

An ROC analysis was used to compare the performance on STAR Math to performance on state achievement tests. The STAR Math Scaled Scores used for analysis originated from assessments 3–11 months before the state achievement test was administered. Selection of cut scores was based on the graph of sensitivity and specificity versus the Scaled Score. For each grade, the Scaled Score chosen as the cut point was equal to the score where sensitivity and specificity intersected. The aggregated and classification analyses, cut points, and outcome measures are outlined in Table 22. When collapsed across ethnicity, AUC values were all greater than 0.80. Descriptive notes for other values represented in the table are provided in the table footnote.

Table 22: Classification Accuracy in Predicting Proficiency on State Achievement Tests in 7 States (Arkansas, Delaware, Illinois, Kansas, Michigan, Mississippi, and North Carolina) and the Terra Nova in Oklahoma

Statistic[a]	Value	
False Positive Rate	0.2559	
False Negative Rate	0.2454	
Sensitivity	0.7546	
Specificity	0.7441	
Positive Predictive Power	0.4683	
Negative Predictive Power	0.9103	
Overall Classification Rate	0.7465	
	Grade	**AUC**
AUC (ROC)	2	0.811
	3	0.820
	4	0.824
	5	0.837
	6	0.852
	7	0.834
	8	0.804
Base Rate	0.23	

Table 22: Classification Accuracy in Predicting Proficiency on State Achievement Tests in 7 States (Arkansas, Delaware, Illinois, Kansas, Michigan, Mississippi, and North Carolina) and the Terra Nova in Oklahoma (Continued)

Statistic[a]	Value	
	Grade	Cut Score
Cut Point	2	421
	3	523
	4	607
	5	658
	6	708
	7	744
	8	759

a. The false positive rate is equal to the proportion of students incorrectly labeled "at-risk." The false negative rate is equal to the proportion of students incorrectly labeled as not "at-risk." Likewise, sensitivity refers to the proportion of correct positive predictions while specificity refers to the proportion of negatives that are correctly identified (e.g. student will *not* meet a particular cut score).

Aggregated Reliability and Validity Data

Tables 23 and 24 provide aggregated reliability values as well as concurrent and predictive validity evidence for STAR Math. All reliability coefficients were greater than 0.90 and median validity coefficients ranged from 0.67–0.80.

Table 23: Overall Reliability Estimates for STAR Math

Type of Reliability	Grade	N	Coefficient	SEM
Generic	Grades 1–12	29,228	0.947	40
Split-Half	Grades 1–12	29,228	0.944	40
Alternate Forms/Test-Retest	Grades 1–12	7,389	0.908	40

Table 24: Overall Concurrent and Predictive Validity Evidence for STAR Math

Type of Validity	Age or Grade	Test or Criterion	N (Range)	Coefficient	
				Range	Median
Concurrent	Grades 3–8	Idaho Standards Achievement Test	2,458 (170–231)	0.68–0.85	0.80
	Grades 3–7	Michigan Educational Assessment Program	1,179 (53–162)	0.58–0.84	0.76
	Grades 2–6, 8	Delaware Student Testing Program	1,330 (44–296)	0.56–0.78	0.72
	Grades 3, 5	Minnesota Comprehensive Assessment	340 (81–91)	0.71–0.76	0.74
	Grades 3–6	Mississippi Curriculum Test	442 (52–154)	0.43–0.78	0.74
	Grades 3–8	Mississippi Curriculum Test 2nd Edition	8,017 (572–1,909)	0.65–0.77	0.72
	Grades 2–5	TerraNova	1,314 (119–205)	0.45–0.78	0.71
Predictive	Grades 3–7	Michigan Educational Assessment Program	840 (37–84)	0.63–0.87	0.78
	Grades 3–5	Mississippi Curriculum Test	583 (33–164)	0.51–0.82	0.71
	Grades 3–8	Texas Assessment of Knowledge and Skills	2,397 (135–646)	0.49–0.74	0.70
	Grades 3–8	Mississippi Curriculum Test 2nd Edition	15,774 (2,148–2,977)	0.60–0.72	0.67

VALIDITY

The key concept used to judge a test's usefulness is its validity. Validity is the degree to which a test measures what it claims to measure. Evidence of STAR Math validity takes many forms, including correlations with teacher ratings of their students' math skills, correlations with scores on a wide variety of published tests with established reliability and validity, and correlations with state accountability tests.

Establishing construct validity involves the use of data and other information external to the test instrument itself. For example, the STAR Math test claims to provide an estimate of a child's mathematical achievement level for use in placement. Therefore, demonstration of STAR Math's construct validity rests on the evidence that the test in fact provides such an estimate.

There are a number of ways to demonstrate this. One method includes examining the relationship between students' STAR Math Scaled Scores and their grade levels. Since mathematical ability varies significantly within and across grade levels and improves as a student's grade level increases, STAR Math data should demonstrate these anticipated relationships. Tables 42 and 43 on page 93 show a consistent pattern of grade over grade increases in average STAR Math Scaled Scores. As STAR Math is psychometrically identical with its earlier incarnations, this pattern is consistent with the proposition that the STAR Math test effectively measures the mathematics achievement of students.

Another source of evidence for construct validity is the relationship between students' STAR Math scores and their scores on other measures of mathematics achievement. If it is a valid assessment, the STAR Math test should correlate highly with other accepted procedures and measures that are used to determine mathematics achievement level. Additionally, these scores should be highly related to teachers' assessments of their students' proficiency in mathematics.

In the remainder of this chapter, validity evidence of two kinds will be presented. First, data that demonstrate a strong and positive correlation between STAR Math scores and scores on other standardized tests will be presented. Second, data that show a strong degree of relationship between STAR Math scores and teacher ratings of their students' proficiency in selected math skills will be presented. All evidence supporting the validity of earlier versions of STAR Math applies perforce to all later versions of the program.

Relationship of STAR Math Scores to Scores on Other Tests of Mathematics Achievement

The technical manual for the earliest version of STAR Math listed correlations between scores on that test and those on a number of other standardized measures of math achievement, obtained in 1998 for more than 9,000 students who participated in STAR Math norming for that version of the program. The standardized tests included a variety of well-established instruments, including the California Achievement Test (CAT), the Comprehensive Test of Basic Skills (CTBS), the Iowa Tests of Basic Skills

(ITBS), the Metropolitan Achievement Test (MAT), the Stanford Achievement Test (SAT), and several statewide tests.

During the current norming of STAR Math, scores on other standardized tests were obtained for more than 30,000 additional students. All of the standardized tests listed above were included, plus others such as Northwest Evaluation Association (NWEA) and TerraNova. Scores on state assessments from the following states were also included: Arkansas, Connecticut, Delaware, Florida, Georgia, Kentucky, Idaho, Indiana, Illinois, Maryland, Michigan, Minnesota, Mississippi, New York, North Carolina, Ohio, Oklahoma, Oregon, Pennsylvania, Rhode Island, South Dakota, Texas, Virginia, and Washington. The extent that the STAR Math test correlates with these tests provides support for its construct validity. That is, strong and positive correlations between STAR Math and these other instruments provide support for the claim that STAR Math effectively measures mathematics achievement.

Tables 25–28 present the correlation coefficients between the scores on the STAR Math test and each of the other test instruments for which data were received. Tables 25 and 26 display “concurrent validity” data, that is, correlations between STAR Math norming study test scores and other tests administered within a two-month time period. Tests listed in Tables 25 and 26 were administered between the fall of 2001 and the fall of 2010. Tables 27 and 28 display all other correlations of STAR Math norming tests and external tests; the external test scores were administered at various times prior to spring 2002, and were obtained from student records.

In addition to the concurrent validity estimates provided in Tables 25 and 26, data concerning STAR Math’s predictive validity are available in Tables 27 and 28. Predictive validity provides an estimate of the extent to which scores on the STAR Math test predicted scores on criterion measures given at a later point in time, operationally defined as more than 2 months between the STAR test (predictor) and the criterion test. It provides an estimate of the linear relationship between STAR scores and scores on measures covering a similar academic domain. Predictive correlations are attenuated by time due to the fact that students are gaining skills in the interim between testing occasions, and also by differences between the tests’ content specifications.

Tables 25–28 are presented in two parts. Tables 25 and 27 display validity coefficients for grades 1–6, and Tables 26 and 28 display the validity coefficients for grades 7–12. The bottom of each table presents a grade-by-grade summary, including the total number of students for whom test data were available, the number of validity coefficients for that grade, and the average value of the validity coefficients.

The within-grade average concurrent validity coefficients for grades 1–6 varied from 0.63–0.73, with an overall average of 0.67. The within-grade average concurrent validity for grades 7–12 ranged from 0.53–0.75, with an overall average of 0.70. Predictive validity coefficients ranged from 0.55–0.73 in grades 1–6, with an average of 0.66. In grades 7–12 the predictive validity coefficients ranged from 0.73–0.80, with an average of 0.77. The other validity coefficient within-grade averages (for STAR Math 2.0 with external tests administered prior to spring 2002, Tables 29 and 30) varied from 0.56–0.70; the overall average was 0.63.

Since correlation coefficients are available for many different test editions, forms, and dates of administration, many of the tests have several validity coefficients associated with them. Where test data quality could not be verified, and when sample size was very small, those data were omitted from the tabulations. Correlations were computed separately on tests according to the unique combination of test edition/form and time when testing occurred. Testing data for other standardized tests administered prior to spring 1998 were excluded from the validity analyses.

In general, these correlation coefficients reflect very well on the validity of the STAR Math test as a tool for placement in mathematics. In fact, the correlations are similar in magnitude to the validity coefficients of these measures with each other. These validity results, combined with the supporting evidence of reliability and minimization of SEM estimates for the STAR Math test, provide a quantitative demonstration of how well this innovative instrument in mathematics achievement assessment performs.

Table 25: Concurrent Validity Data: STAR Math 2.x Correlations (r) with External Tests Administered Spring 2002–Spring 2010, Grades 1–6[a]

Test Form	Date	Score	1		2		3		4		5		6	
			n	r	n	r	n	r	n	r	n	r	n	r
Arkansas Augmented Benchmark Examination (AABE)														
AABE	S 08	SS	–	–	–	–	725	0.68*	686	0.70*	634	0.70*	297	0.66*
California Achievement Test (CAT) 5th Edition														
CAT	S 02	NCE	–	–	–	–	17	0.50*	–	–	–	–	–	–
Comprehensive Test of Basic Skills (CTBS)														
CTBS–A13	S 02	SS	–	–	–	–	–	–	–	–	21	0.66*	–	–
CTBS	S 02	NCE	–	–	–	–	–	–	–	–	–	–	32	0.65*
Delaware Student Testing Program (DSTP)														
DSTP	S 03	SS	–	–	–	–	258	0.72*	–	–	296	0.73*	–	–
DSTP	S 05	SS	–	–	–	–	66	0.67*	–	–	–	–	–	–
DSTP	S 06	SS	–	–	140	0.66*	58	0.85*	40	0.63*	151	0.75*	44	0.77*
Florida Comprehensive Assessment Test (FCAT)														
FCAT	S 06	SS	–	–	–	–	58	0.85*	40	0.63*	–	–	–	–
FCAT	S 06–08	SS	–	–	–	–	2,338	0.74*	2,211	0.74*	2,078	0.74*	279	0.65*
Idaho Standards Achievement Test (ISAT)														
ISAT	F 02	SS	–	–	–	–	192	0.68*	188	0.75*	194	0.75*	221	0.74*
ISAT	S 03	SS	–	–	–	–	224	0.74*	209	0.83*	222	0.78*	231	0.82*
ISAT	S 07–09	SS	–	–	–	–	798	0.70*	699	0.60*	727	0.62*	217	0.69*

Table 25: Concurrent Validity Data: STAR Math 2.x Correlations (r) with External Tests Administered Spring 2002–Spring 2010, Grades 1–6[a] (Continued)

Test Form	Date	Score	1		2		3		4		5		6	
			n	r	n	r	n	r	n	r	n	r	n	r
Iowa Test of Basic Skills (ITBS)														
ITBS–A	S 02	NCE	–	–	–	–	–	–	50	0.66*	79	0.72*	–	–
ITBS–K	S 02	SS	–	–	–	–	–	–	–	–	–	–	70	0.69*
ITBS–L	S 02	NCE	–	–	7	0.78*	23	0.57*	17	0.70*	21	0.66*	–	–
ITBS–M	S 02	NCE	14	0.56*	11	0.58	–	–	–	–	–	–	–	–
ITBS–M	S 02	SS	–	–	–	–	17	0.72*	–	–	–	–	–	–
Kansas State Assessment Program (KSAP)														
KSAP	S 06–08	SS	–	–	–	–	915	0.59*	947	0.67*	752	0.66*	402	0.67*
Kentucky Core Content Test (KCCT)														
KCCT	S 08–10	SS	–	–	–	–	3,777	0.69*	3,115	0.70*	2,228	0.66*	1,785	0.66*
McGraw Hill Mississippi/Criterion Referenced														
–	S 02	SS	–	–	–	–	–	–	–	–	44	0.73*	–	–
Metropolitan Achievement Test (MAT)														
MAT–6th Ed.	S 02	NCE	69	0.55*	–	–	–	–	–	–	–	–	–	–
MAT–8th Ed.	S 02	SS	–	–	–	–	–	–	38	0.83	–	–	–	–
Michigan Educational Assessment Program (MEAP) – Mathematics														
MEAP	F 04	SS	–	–	–	–	–	–	154	0.81*	–	–	–	–
MEAP	F 05	SS	–	–	–	–	71	0.75*	69	0.78*	77	0.83*	89	0.77*
MEAP	F 06	SS	–	–	–	–	162	0.72*	–	–	53	0.67*	123	0.69*
Minnesota Comprehensive Assessment (MCA)														
MCA	S 03	SS	–	–	–	–	85	0.71*	–	–	81	0.76*	–	–
MCA	S 04	SS	–	–	–	–	91	0.74*	–	–	83	0.73*	–	–
Mississippi Curriculum Test (MCT2)														
CTB	S 02	SS	–	–	–	–	–	–	10	0.62	–	–	–	–
CTB	S 03	SS	–	–	–	–	117	0.71*	154	0.77*	119	0.78*	52	0.43*
MCT	S 03	SS	–	–	–	–	117	0.71*	154	0.77*	110	0.78*	52	0.43
MCT2	S 08	SS	–	–	–	–	1,786	0.72*	1,757	0.72*	1,531	0.73*	1,180	0.78*

Table 25: Concurrent Validity Data: STAR Math 2.x Correlations (r) with External Tests Administered Spring 2002–Spring 2010, Grades 1–6[a] (Continued)

Test Form	Date	Score	1		2		3		4		5		6	
			n	r	n	r	n	r	n	r	n	r	n	r
North Carolina End-of-Grade (NCEOG) Test														
NCEOG	S 02	NCE	–	–	–	–	70	0.60*						
NCEOG	S 02	SS					62	0.73*						
NCEOG	S 06–08	SS	–	–	–	–	1,100	0.72*	751	0.72*	482	0.65*	202	0.77*
NWEA, NALT, & MAP														
	F 02	SS	–	–	–	–	81	0.75*	–	–	77	0.86*	–	–
	S 03	SS	–	–	–	–	85	0.82*	–	–	80	0.85*	–	–
	F 03	SS	–	–	77	0.69*	92	0.73*	75	0.82*	79	0.86*	–	–
	S 04	SS	–	–	80	0.72*	92	0.84*	65	0.84*	82	0.86*	–	–
	F 04	SS	–	–	–	–	63	0.53*	77	0.78*	86	0.84*	–	–
	S 05	SS	–	–	–	–	63	0.74*	80	0.87*	96	0.87*	–	–
Oklahoma Core Curriculum Test (OCCT)														
OCCT	S 06	SS	–	–	–	–	77	0.71*	92	0.61*	66	0.68*	60	0.63*
Pennsylvania System of School Assessment (PSSA)														
PSSA	S 02	SS	–	–	–	–	–	–	–	–	–	–	62	0.76*
South Dakota State Test of Educational Progress (DSTEP)														
DSTEP	S 08–10	SS	–	–	–	–	2,092	0.74*	1,555	0.74*	1,309	0.72*	837	0.74*
Stanford Achievement Test (SAT9)														
SAT9	S 02	NCE	–	–	113	0.56*	39	0.83*	46	0.54*	103	0.70*	49	0.65
SAT9	S 02	SS	20	0.76*	16	0.68*	18	0.59*	19	0.57*	71	0.49*	84	0.62*
TerraNova														
TerraNova	S 02	NCE	7	0.66	14	0.46	125	0.68*	18	0.67*	17	0.79*	15	0.64
TerraNova	F 03	SS	–	–	177	0.55*	172	0.45*	119	0.67*	160	0.78*	–	–
TerraNova	S 04	SS	–	–	150	0.75*	205	0.71*	149	0.71*	182	0.78*	–	–
Texas Assessment of Academic Achievement (TAAS)														
TAAS	S 01	SS	–	–	–	–	1,036	0.56*	1,047	0.50*	1,006	0.65*	991	0.61*
TAAS	S 02	SS	–	–	–	–	674	0.65*	669	0.63*	677	0.64*	885	0.64*

Table 25: Concurrent Validity Data: STAR Math 2.x Correlations (r) with External Tests Administered Spring 2002–Spring 2010, Grades 1–6[a] (Continued)

Test Form	Date	Score	1		2		3		4		5		6	
			n	r	n	r	n	r	n	r	n	r	n	r
Texas Assessment of Knowledge and Skills (TAKS)														
TAKS	S 03	SS	–	–	–	–	1,134	0.63*	1,129	0.62*	1,086	0.70*	–	–
Wisconsin Knowledge and Concepts Examination (WKCE)														
WKCE	F 06–10	SS	–	–	–	–	1,322	0.71*	1,393	0.72*	1,801	0.73*	1,175	0.75*
Summary														
Grade(s)	**All**		**1**		**2**		**3**		**4**		**5**		**6**	
Number of students	65,609		110		785		20,497		17,822		16,961		9,434	
Number of coefficients	152		4		10		41		34		38		25	
Average validity	–		0.63		0.64		0.69		0.70		0.73		0.68	
Overall average	0.67													

a. Asterisk (*) denotes correlation coefficients that are statistically significant at the 0.05 level.

Table 26: Concurrent Validity Data: STAR Math 2.x Correlations (r) with External Tests Administered Spring 2002–Spring 2010, Grades 7–12[a]

Test Form	Date	Score	7		8		9		10		11		12	
			n	r	n	r	n	r	n	r	n	r	n	r
Arkansas Augmented Benchmark Examination (AABE)														
AABE	S 08	SS	99	0.56*	74	0.77*	–	–	–	–	–	–	–	–
Delaware Student Testing Program (DSTP)														
DSTP	S 03	SS	–	–	254	0.78*	–	–	–	–	–	–	–	–
Florida Comprehensive Assessment Test (FCAT)														
FCAT	S 02	SS	–	–	–	–	–	–	51	0.64*	57	0.66*	38	0.75*
FCAT	S 06–08	SS	195	0.65*	89	0.60*	–	–	–	–	–	–	–	–
Idaho Standards Achievement Test (ISAT)														
ISAT	F 02	SS	206	0.81*	170	0.81*	–	–	–	–	–	–	–	–
ISAT	S 03	SS	227	0.85*	174	0.82*	–	–	–	–	–	–	–	–
ISAT	S 06–08	SS	289	0.71*	328	0.77*	–	–	–	–	–	–	–	–

Table 26: Concurrent Validity Data: STAR Math 2.x Correlations (r) with External Tests Administered Spring 2002–Spring 2010, Grades 7–12[a] (Continued)

Test Form	Date	Score	7		8		9		10		11		12	
			n	r	n	r	n	r	n	r	n	r	n	r
Iowa Test of Basic Skills (ITBS)														
ITBS–M	S 02	SS	37	0.40*	–	–	–	–	–	–	–	–	–	–
Kansas State Assessment Program (KSAP)														
KSAP	S 06–08	SS	271	0.74*	137	0.75*	–	–	–	–	–	–	–	–
Kentucky Core Content Test (KCCT)														
KCCT	S 08–10	SS	788	0.68*	362	0.64*	–	–	–	–	–	–	–	–
Michigan Educational Assessment Program (MEAP) – Mathematics														
MEAP	F 05	SS	65	0.72*	71	0.80*	–	–	–	–	–	–	–	–
MEAP	F 06	SS	122	0.84*	123	0.58*	–	–	–	–	–	–	–	–
Mississippi Curriculum Test (MCT2)														
MCT2	S 08	SS	721	0.66*	549	0.71*	–	–	–	–	–	–	–	–
New Standards Reference Mathematics Exam (Rhode Island)														
NRSME	S 02	SS	–	–	–	–	–	–	–	–	67	0.67*	9	0.66
North Carolina End-of-Grade (NCEOG) Test														
NCEOG	S 06–08	SS	216	0.70*	39	0.81*	–	–	–	–	–	–	–	–
Ohio Proficiency Test (OPT)														
OPT	S 02	SS	–	–	–	–	23	0.67*	26	0.40*	24	0.77*	24	0.69*
Oklahoma Core Curriculum Test (OCCT)														
OCCT	S 06	SS	55	0.63*	68	0.70*	–	–	–	–	–	–	–	–
Otis Lennon School Ability Test (OLSAT)														
OLSAT	S 02	NCE	–	–	–	–	–	–	12	0.36	13	0.91*	6	0.72
Palmetto Achievement Challenge Test (PACT), 2001														
PACT	S 02	SS	–	–	161	0.72*	–	–	–	–	–	–	–	–
South Dakota State Test of Educational Progress (DSTEP)														
DSTEP	S 08–10	SS	525	0.73*	535	0.73*	–	–	–	–	–	–	–	–
Texas Assessment of Academic Achievement (TAAS)														
TAAS	S 01	SS	892	0.60*	825	0.67*	–	–	–	–	–	–	–	–
TAAS	S 02	SS	768	0.62*	809	0.68*	–	–	–	–	–	–	–	–
Texas Assessment of Academic Skills (TAAS), 2001														
TAAS	S 02	TLI	–	–	–	–	163	0.69*	–	–	–	–	–	–

Table 26: Concurrent Validity Data: STAR Math 2.x Correlations (r) with External Tests Administered Spring 2002–Spring 2010, Grades 7–12[a] (Continued)

Test Form	Date	Score	7		8		9		10		11		12	
			n	r	n	r	n	r	n	r	n	r	n	r
Wisconsin Knowledge and Concepts Examination (WKCE)														
WKCE	F 06–10	SS	640	0.79*	767	0.76*	–	–	248	0.73*	–	–	–	–
Summary														
Grade(s)	**All**		**7**		**8**		**9**		**10**		**11**		**12**	
Number of students	12,412		6,116		5,535		186		337		161		77	
Number of coefficients	49		17		18		2		4		4		4	
Average validity	–		0.69		0.73		0.68		0.53		0.75		0.71	
Overall average	0.70													

a. Asterisk (*) denotes correlation coefficients that are statistically significant at the 0.05 level.

Table 27: Predictive Validity Data: STAR Math 2.x Correlations (r) with External Tests Administered Fall 2001–Fall 2009, Grades 1–6[a]

Test Form	Date[b]	Score	1		2		3		4		5		6	
			n	r	n	r	n	r	n	r	n	r	n	r
Arkansas Augmented Benchmark Examination (AABE)														
AABE	F 07	SS	–	–	–	–	1,196	0.69*	1,128	0.67*	994	0.73*	638	0.71*
Delaware Student Testing Program (DSTP)														
DSTP	F 02	SS	–	–	–	–	191	0.70*	–	–	228	0.70*	–	–
DSTP	F 04	SS	–	–	–	–	171	0.67*	–	–	–	–	–	–
DSTP	W 05	SS	–	–	–	–	149	0.76*	–	–	–	–	–	–
DSTP	S 05	SS	–	–	–	–	132	0.64*	172	0.63*	185	0.62*	–	–
DSTP	F 05	SS	–	–	206	0.64*	219	0.66*	249	0.67*	265	0.68*	–	–
DSTP	W 05	SS	–	–	242	0.61*	226	0.61*	269	0.62	277	0.68		
Florida Comprehensive Assessment Test (FCAT)														
FCAT	F 05	SS	–	–	–	–	54	0.79*	42	0.69*	–	–	–	–
FCAT	F 05–07	SS	–	–	–	–	5,292	0.74*	5,020	0.73*	4,895	0.77*	1,015	0.66*
Idaho Standards Achievement Test (ISAT)														
ISAT	F 08–10	SS	–	–	–	–	1,875	0.67*	1,908	0.63*	2,312	0.69*	1,809	0.73*

Table 27: Predictive Validity Data: STAR Math 2.x Correlations (r) with External Tests Administered Fall 2001–Fall 2009, Grades 1–6[a] (Continued)

Test Form	Date[b]	Score	1		2		3		4		5		6	
			n	r	n	r	n	r	n	r	n	r	n	r
Kentucky Core Content Test (KCCT)														
KCCT	F 07–09	SS	–	–	–	–	5,821	0.68*	5,325	0.67*	4,199	0.66*	3,172	0.63*
Michigan Educational Assessment Program (MEAP)														
MEAP	F 04	SS	–	–	–	–	–	–	64	0.70*	74	0.85*	81	0.74*
MEAP	W 05	SS	–	–	–	–	–	–	65	0.80*	75	0.87*	42	0.72*
MEAP	S 05	SS	–	–	–	–	66	0.63*	65	0.73*	76	0.83*	84	0.71*
Minnesota Comprehensive Assessment (MCA)														
MCA	F 02	SS	–	–	–	–	81	0.64*	–	–	78	0.72*	–	–
MCA	W 03	SS	–	–	–	–	86	0.66*	–	–	81	0.77*	–	–
MCA	F 03	SS	–	–	–	–	87	0.53*	–	–	79	0.69*	–	–
MCA	W 04	SS	–	–	–	–	93	0.60*	–	–	82	0.75	–	–
Mississippi Curriculum Test (MCT2)														
MCT	F 02	SS	–	–	–	–	48	0.64*	33	0.82*	73	0.80*	–	–
MCT	F 03	SS	–	–	–	–	109	0.51*	164	0.72*	156	0.69*	–	–
MCT2	F 07	SS	–	–	–	–	2,989	0.69*	3,022	0.70*	2,796	0.72*	2,741	0.74*
North Carolina End-of-Grade (NCEOG) Test														
NCEOG	F 05–07	SS	–	–	–	–	2,494	0.73*	2,008	0.70*	1,096	0.69*	830	0.70*
NWEA NALT & MAP														
	F 02	–	–	–	–	–	80	0.65*	–	–	77	0.86*	–	–
	W 03	–	–	–	–	–	85	0.78*	–	–	80	0.90*	–	–
	F 03	–	–	–	–	–	86	0.68*	69	0.81*	78	0.87*	–	–
	W 04	–	–	–	–	–	92	0.80*	68	0.80*	81	0.93*	–	–
Oklahoma Core Curriculum Test (OCCT)														
OCCT	F 05	SS	–	–	–	–	87	0.71*	88	0.61*	77	0.55*	83	0.56*
South Dakota State Test of Educational Progress (DSTEP)														
DSTEP	F 07–09	SS	–	–	–	–	3,886	0.73*	3,665	0.75*	3,084	0.72*	2,328	0.75*

Table 27: Predictive Validity Data: STAR Math 2.x Correlations (r) with External Tests Administered Fall 2001–Fall 2009, Grades 1–6[a] (Continued)

Test Form	Date[b]	Score	1		2		3		4		5		6	
			n	r	n	r	n	r	n	r	n	r	n	r
STAR Math														
STAR–M	F 01	SS	–	–	–	–	1,036	0.61*	1,047	0.63*	1,006	0.65*	991	0.65*
STAR–M	F 05	SS	2,605	0.50*	7,195	0.63*	11,716	0.67*	13,295	0.69*	10,343	0.70*	6,823	0.75*
STAR–M	F 06	SS	4,687	0.58*	12,464	0.62*	16,474	0.66*	17,161	0.70*	16,181	0.71*	12,026	0.73*
STAR–M	F 05	SS	1,147	0.51*	3,181	0.62*	4,894	0.67*	5,254	0.70*	2,164	0.69*	1,474	0.74*
STAR–M	F 05	SS	1,147	0.42*	3,181	0.57*	4,894	0.62*	5,254	0.64*	2,164	0.73*	1,474	0.80*
STAR–M	S 06	SS	1,147	0.66*	3,181	0.69*	4,894	0.73*	5,254	0.74*	2,164	0.73*	1,474	0.80*
STAR–M	S 06	SS	1,147	0.62*	3,181	0.63*	4,894	0.69*	5,254	0.70*	2,164	0.71*	1,474	0.78*
Texas Assessment of Academic Achievement (TAAS)														
TAAS	F 01	SS	–	–	–	–	1,036	0.51*	1,047	0.42*	1,006	0.60*	991	0.61*
Texas Assessment of Knowledge and Skills (TAKS)														
TAKS	F 02	SS	–	–	–	–	262	0.64*	135	0.49*	228	0.70*	646	0.69*
TerraNova														
TerraNova	F 03	–	–	–	117	0.69*	165	0.58*	116	0.75*	154	0.54*	–	–
TerraNova	W 04	–	–	–	128	0.58*	197	0.47*	120	0.71*	173	0.77*	–	–
Wisconsin Knowledge and Concepts Examination (WKCE)														
WKCE	S 05–09	SS	–	–	–	–	4,645	0.66*	4,980	0.68	5,345	0.74*	4,702	0.75*

Summary

Grade(s)	All	1	2	3	4	5	6
Number of students	317,587	11,880	33,076	80,802	82,341	64,590	44,898
Number of coefficients	143	6	10	38	31	37	21
Average validity	–	0.55	0.63	0.66	0.69	0.73	0.71
Overall average	0.66						

a. Asterisk (*) denotes correlation coefficients that are statistically significant at the 0.05 level.

b. Dates correspond to the term and year of the predicting scores.

Table 28: Predictive Validity Data: STAR Math 2.x Correlations (r) with External Tests Administered Fall 2001–Fall 2009, Grades 7–12[a]

Test Form	Date[b]	Score	7		8		9		10		11		12	
			n	r	n	r	n	r	n	r	n	r	n	r
Arkansas Augmented Benchmark Examination (AABE)														
AABE	F 07	SS	369	0.67*	296	0.76*	–	–	–	–	–	–	–	–
Delaware Student Testing Program (DSTP)														
DSTP	F 02	SS	242	0.74*	–	–	–	–	–	–	–	–	–	–
DSTP	S 05	SS	227	0.71*	175	0.75*	–	–	–	–	–	–	–	–
Florida Comprehensive Assessment Test (FCAT)														
FCAT	F 05–07	SS	783	0.72*	336	0.70*	–	–	–	–	–	–	–	–
Idaho Standards Achievement Test (ISAT)														
ISAT	F 05–07	SS	588	0.75*	484	0.75*	–	–	–	–	–	–	–	–
Kentucky Core Content Test (KCCT)														
KCCT	F 07–09	SS	1,789	0.65*	1,153	0.59*	–	–	–	–	–	–	–	–
Michigan Educational Assessment Program (MEAP)														
MEAP	F 04	SS	56	0.78*	–	–	–	–	–	–	–	–	–	–
MEAP	W 05	SS	56	0.78*	–	–	–	–	–	–	–	–	–	–
MEAP	S 05	SS	37	0.86*	–	–	–	–	–	–	–	–	–	–
Mississippi Curriculum Test (MCT2)														
MCT2	F 07	SS	2,127	0.71*	2,190	0.70*	–	–	–	–	–	–	–	–
North Carolina End-of-Grade (NCEOG) Test														
NCEOG	F 05–07	SS	443	0.78*	397	0.71*	–	–	–	–	–	–	–	–
Oklahoma Core Curriculum Test (OCCT)														
OCCT	F 05	SS	74	0.57*	70	0.67*	–	–	–	–	–	–	–	–
South Dakota State Test of Educational Progress (DSTEP)														
DSTEP	F 07–09	SS	1,851	0.74*	1,522	0.75*	–	–	–	–	–	–	–	–

Table 28: Predictive Validity Data: STAR Math 2.x Correlations (r) with External Tests Administered Fall 2001–Fall 2009, Grades 7–12[a] (Continued)

Test Form	Date[b]	Score	7		8		9		10		11		12	
			n	r	n	r	n	r	n	r	n	r	n	r
STAR Math														
STAR–M	F 01	–	892	0.72*	825	0.78*	–	–	–	–	–	–	–	–
STAR–M	F 05	–	3,551	0.75*	2,693	0.76*	668	0.79*	508	0.79*	572	0.79*	378	0.76*
STAR–M	F 06	–	7,564	0.76*	7,122	0.77*	1,017	0.78*	876	0.76*	693	0.83*	507	0.77*
STAR–M	F 05	–	1,191	0.75*	127	0.84*	215	0.78*	213	0.83*	164	0.75*	–	–
STAR–M	F 05	–	1,191	0.71*	127	0.77*	215	0.78*	213	0.81*	164	0.75*	–	–
STAR–M	S 06	–	1,191	0.79*	127	0.82*	215	0.80*	213	0.85*	164	0.79*	–	–
STAR–M	S 06	–	1,191	0.77*	127	0.82*	215	0.76*	213	0.82*	164	0.77*	–	–
Texas Assessment of Academic Achievement (TAAS)														
TAAS	F 01	SS	892	0.59*	825	0.67*	–	–	–	–	–	–	–	–
Texas Assessment of Knowledge and Skills (TAKS)														
TAKS	F 02	SS	564	0.74*	562	0.74*	–	–	–	–	–	–	–	–
Wisconsin Knowledge and Concepts Examination (WKCE)														
WKCE	S 05–09	SS	1,883	0.79*	1,742	0.76*	–	–	289	0.76*	–	–	–	–

Summary

Grade(s)	All	7	8	9	10	11	12
Number of students	57, 528	28,752	20,900	2,545	2,525	1,921	885
Number of coefficients	63	23	19	6	7	6	2
Average validity	–	0.73	0.74	0.78	0.80	0.78	0.77
Overall average	0.77						

a. Asterisk (*) denotes correlation coefficients that are statistically significant at the 0.05 level.

b. Dates correspond to the term and year of the predicting scores.

Table 29: Other External Validity Data—STAR Math 2.0 Correlation Coefficients (r) with External Tests Administered Prior to Spring 2002, Grades 1–6[a]

Test	Version	Date	Score	1		2		3		4		5		6	
				n	r	n	r	n	r	n	r	n	r	n	r
Achievement Level (RIT) Test															
RIT		F 01	SS	–	–	–	–	–	–	–	–	–	–	150	0.69*
California Achievement Test															
CAT	5th Ed.	S 01	SS	–	–	–	–	46	0.52*	–	–	–	–	–	–
Cognitive Abilities Test															
CogAT		F 00	SS	–	–	–	–	41	0.61*	–	–	–	–	–	–
CogAT		F 01	SS	–	–	45	0.73*	–	–	–	–	–	–	–	–
Comprehensive Test of Basic Skills															
CTBS	4th Ed.	S 01	GE	–	–	–	–	–	–	43	0.67*	–	–	–	–
CTBS	A-13	S 00	NCE	–	–	–	–	–	–	65	0.60*	–	–	–	–
CTBS	A-13	S 00	SS	–	–	–	–	–	–	–	–	44	0.70*	–	–
CTBS	A-13	S 01	GE	–	–	–	–	–	–	–	–	–	–	56	0.69*
CTBS	A-13	S 01	NCE	–	–	–	–	–	–	–	–	67	0.72*	–	–
CTBS	A-13	S 01	SS	–	–	–	–	–	–	42	0.61*	–	–	–	–
Connecticut Mastery Test															
Conn	2nd	F 00	SS	–	–	–	–	–	–	–	–	35	0.51*	–	–
Conn	3rd	F 01	SS	–	–	–	–	–	–	42	0.64*	–	–	27	0.52*
Des Moines Public School (Grade 2 pretest)															
DMPS		F 01	NCE	–	–	25	0.76*	–	–	–	–	–	–	–	–
Educational Development Series															
EDS	13C	S 01	GE	–	–	–	–	30	0.69*	–	–	–	–	–	–
EDS	14C	S 00	GE	–	–	–	–	–	–	32	0.44*	–	–	–	–
EDS	15C	F 01	GE	–	–	–	–	–	–	–	–	37	0.68*	–	–
Florida Comprehensive Assessment Test															
FCAT		S 01	NCE	–	–	–	–	–	–	–	–	73	0.65*	–	–

Table 29: Other External Validity Data—STAR Math 2.0 Correlation Coefficients (r) with External Tests Administered Prior to Spring 2002, Grades 1–6[a] (Continued)

Test	Version	Date	Score	1		2		3		4		5		6	
				n	r	n	r	n	r	n	r	n	r	n	r
Iowa Tests of Basic Skills															
ITBS	Form A	S 01	NCE	–	–	–	–	73	0.45*	78	0.65*	–	–	–	–
ITBS	Form A	F 01	NCE	–	–	–	–	25	0.41*	25	0.35	23	0.33	86	0.81*
ITBS	Form A	F 01	SS	–	–	–	–	–	–	–	–	–	–	73	0.64*
ITBS	Form K	F 00	SS	–	–	–	–	–	–	–	–	–	–	20	0.92*
ITBS	Form K	S 01	NCE	–	–	101	0.67*	74	0.64*	31	0.25	11	0.58	31	0.62*
ITBS	Form K	F 01	NCE	–	–	–	–	10	0.78*	16	0.78*	9	0.54	18	0.63*
ITBS	Form K	F 01	SS	–	–	–	–	–	–	–	–	75	0.77*	68	0.71*
ITBS	Form L	S 01	NCE	–	–	–	–	13	0.50	46	0.81*	13	0.73*	–	–
ITBS	Form L	S 01	SS	–	–	–	–	–	–	11	0.81*	–	–	–	–
ITBS	Form L	F 01	NCE	–	–	–	–	–	–	–	–	69	0.66*	–	–
ITBS	Form M	S 99	NCE	–	–	–	–	–	–	–	–	–	–	19	0.68*
ITBS	Form M	S 00	NCE	–	–	–	–	–	–	–	–	28	0.65*	–	–
ITBS	Form M	S 01	NCE	–	–	19	0.81*	–	–	43	0.78*	–	–	–	–
ITBS	Form M	S 01	SS	–	–	–	–	47	0.39*	32	0.55*	–	–	–	–
ITBS	Form M	F 01	NCE	5	0.88*	–	–	–	–	15	0.82*	–	–	–	–
McGraw Hill Mississippi/Criterion Referenced															
McGraw		S 01	SS	–	–	–	–	–	–	–	–	121	0.52*	–	–
Metropolitan Achievement Test															
MAT	7th Ed.	F 01	NCE	–	–	–	–	–	–	–	–	–	–	15	0.84*
Michigan Education Assessment Program															
MEAP		S 01	SS	–	–	–	–	–	–	–	–	88	0.72*	–	–
Multiple Assessment Series (Primary Grades)															
Multiple		S 01	NCE	–	–	14	0.52	19	0.54*	–	–	–	–	–	–
New York State Math Assessment															
NYSMA		S 01	SS	–	–	–	–	–	–	–	–	50	0.79*	–	–
North Carolina End of Grade															
NCEOG		F 01	SS	–	–	–	–	85	0.57*	–	–	–	–	–	–

Table 29: Other External Validity Data—STAR Math 2.0 Correlation Coefficients (r) with External Tests Administered Prior to Spring 2002, Grades 1–6[a] (Continued)

Test	Version	Date	Score	1		2		3		4		5		6	
				n	r	n	r	n	r	n	r	n	r	n	r
Northwest Evaluation Association Levels Test															
NWEA		S 01	NCE	–	–	–	–	–	–	–	–	83	0.81*	64	0.78*
NWEA		F 01	NCE	–	–	–	–	50	0.56*	49	0.54*	99	0.70*	–	–
Ohio Proficiency Test															
Ohio		S 01	SS	–	–	–	–	113	0.65*	–	–	–	–	–	–
Stanford Achievement Test															
SAT9		S 99	SS	–	–	–	–	–	–	–	–	55	0.65*	–	–
SAT9		S 00	SS	–	–	–	–	–	–	–	–	–	–	15	0.50
SAT9		F 00	NCE	–	–	–	–	17	0.84*	20	0.83*	–	–	–	–
SAT9		F 00	SS	–	–	–	–	–	–	–	–	–	–	46	0.58*
SAT9		S 01	NCE	–	–	–	–	43	0.69*	–	–	50	0.38*	–	–
SAT9		S 01	SS	64	0.52*	–	–	–	–	58	0.41*	52	0.58*	51	0.65*
SAT9		F 01	SS	–	–	–	–	–	–	90	0.54*	32	0.67*	24	0.57*
Tennessee Comprehensive Assessment Program, 2001															
TCAP	2001	S 01	SS	–	–	–	–	–	–	–	–	48	0.56*	–	–
TerraNova															
TerraNova		S 00	NCE	–	–	–	–	–	–	–	–	–	–	43	0.60*
TerraNova		S 00	SS	–	–	–	–	–	–	–	–	11	0.61*	–	–
TerraNova		F 00	SS	–	–	–	–	–	–	–	–	108	0.62*	–	–
TerraNova		S 01	NCE	–	–	–	–	–	–	–	–	69	0.40*	85	0.62*
TerraNova		S 01	SS	–	–	–	–	–	–	104	0.50*	62	0.59*	131	0.71*
TerraNova		F 01	NCE	–	–	58	0.38*	63	0.56*	70	0.74*	85	0.61*	–	–
Test of New York State Standards															
TONYSS		S 01	SS	–	–	–	–	55	0.75*	68	0.47*	–	–	–	–
Texas Assessment of Academic Skills															
TAAS	2001	S 01	SS	–	–	–	–	–	–	78	0.52*	–	–	–	–
TAAS	2001	S 01	TLI	–	–	–	–	–	–	–	–	–	–	82	0.42*
Virginia Standards of Learning															
Virginia		S 00	SS	–	–	–	–	–	–	–	–	24	0.73*	–	–

Table 29: Other External Validity Data—STAR Math 2.0 Correlation Coefficients (r) with External Tests Administered Prior to Spring 2002, Grades 1–6[a] (Continued)

Test	Version	Date	Score	1		2		3		4		5		6	
				n	r	n	r	n	r	n	r	n	r	n	r
Washington Assessment of Student Learning															
Wash		S 00	SS	–	–	–	–	–	–	–	–	–	–	90	0.54*
Wide Range Achievement Test															
WRAT III		F 01	NCE	–	–	–	–	–	–	44	0.32*	44	0.66*	–	–

Summary

Grade(s)	All	1	2	3	4	5	6
Number of students	4,996	69	262	804	1,102	1,565	1,194
Number of coefficients	98	2	6	17	23	29	21
Average validity	–	0.70	0.65	0.60	0.59	0.62	0.65
Overall average	0.62						

a. n = Sample size.

* Denote correlation coefficients that are statistically significant at the 0.05 level.

Table 30: Other External Validity Data—STAR Math 2.0 Correlation Coefficients (r) with External Tests Administered Prior to Spring 2002, Grades 7–12[a]

Test	Version	Date	Score	7		8		9		10		11		12	
				n	r	n	r	n	r	n	r	n	r	n	r
American College Testing Program															
ACT		F 01	NCE	–	–	–	–	–	–	–	–	–	–	26	0.87*
California Achievement Tests															
CAT	5th Ed.	F 01	NCE	–	–	–	–	64	0.73*	–	–	–	–	–	–
CAT	5th Ed.	F 01	SS	170	0.54*	–	–	–	–	–	–	–	–	–	–
Comprehensive Test of Basic Skills															
CTBS	4th Ed.	S 00	SS	67	0.67*	75	0.73*	–	–	–	–	–	–	–	–
CTBS	A-13	S 00	SS	–	–	31	0.65*	–	–	–	–	–	–	–	–
CTBS	A-13	S 01	SS	23	0.82*	–	–	–	–	48	0.63*	–	–	–	–
Delaware Student Testing Program															
DSTP		S 01	SS	–	–	–	–	94	0.27*	–	–	–	–	–	–
Differential Aptitude Tests															
DAT	Level 1	F 01	NCE	–	–	–	–	41	0.70*	–	–	–	–	–	–

Table 30: Other External Validity Data—STAR Math 2.0 Correlation Coefficients (r) with External Tests Administered Prior to Spring 2002, Grades 7–12[a] (Continued)

Test	Version	Date	Score	7		8		9		10		11		12	
				n	r	n	r	n	r	n	r	n	r	n	r
Explore Tests															
Explore		F 01	NCE	–	–	64	0.54*	–	–	–	–	–	–	–	–
Georgia High School Graduation Test															
Georgia		S 01	NCE	–	–	–	–	–	–	–	–	–	–	23	0.71*
Indiana Statewide Testing for Educational Progress															
ISTEP		F01	NCE	–	–	–	–	51	0.57*	22	0.58*	–	–	–	–
Iowa Tests of Basic Skills															
ITBS	Form A	F 01	SS	66	0.71*	–	–	–	–	–	–	–	–	–	–
ITBS	Form K	S 01	NCE	73	0.80*	18	0.52*	–	–	–	–	–	–	–	–
ITBS	Form K	F 01	NCE	6	0.72	14	0.69*	–	–	–	–	–	–	–	–
ITBS	Form L	S 01	NCE	36	0.74*	32	0.53*	–	–	19	0.67*	32	0.84*	–	–
ITBS	Form M	S 99	NCE	–	–	5	0.89*	–	–	–	–	11	0.80*	–	–
ITBS	Form M	S 00	NCE	–	–	–	–	–	–	9	0.94*	–	–	–	–
ITBS	Form M	S 01	NCE	49	0.52*	48	0.51*	–	–	–	–	–	–	–	–
Kentucky Core Content Test															
KCCT		S 01	NCE	–	–	–	–	45	0.43*	–	–	–	–	–	–
Maryland High School Placement Test															
Maryland		S 01	NCE	–	–	–	–	47	0.60*	–	–	–	–	–	–
McGraw Hill Mississippi/Criterion Referenced															
McGraw		S 01	SS	–	–	–	–	73	0.56*	–	–	–	–	–	–
Metropolitan Achievement Test															
MAT	7th Ed.	F 01	NCE	5	0.80	11	0.82*	–	–	–	–	–	–	–	–
North Carolina End of Grade Tests															
NCEOG		S 01	SS	–	–	177	0.59*	–	–	–	–	–	–	–	–
Oklahoma School Testing Program Core Curriculum Tests															
Oklahoma		S 01	SS	–	–	–	–	26	0.67*	–	–	–	–	–	–

Table 30: Other External Validity Data—STAR Math 2.0 Correlation Coefficients (r) with External Tests Administered Prior to Spring 2002, Grades 7–12[a] (Continued)

Test	Version	Date	Score	7		8		9		10		11		12	
				n	r	n	r	n	r	n	r	n	r	n	r
Oregon State Assessment															
Oregon		S 01	NCE	–	–	45	0.53*	–	–	–	–	–	–	–	–
PLAN															
PLAN		F 99	SS	–	–	–	–	–	–	–	–	–	–	–	0.42
PLAN		F 00	SS	–	–	–	–	–	–	–	–	40	0.28	–	–
PLAN		F 01	NCE	–	–	–	–	–	–	63	0.61*	–	–	–	–
Preliminary SAT/National Merit Scholarship Qualifying Test															
PSAT/NMSQT	NMSQT	F 00	NCE	–	–	–	–	–	–	–	–	–	–	–	0.63*
PSAT/NMSQT	NMSQT	F 01	NCE	–	–	–	–	–	–	–	–	72	0.64*	–	–
Stanford Achievement Test															
SAT9		S 98	NCE	11	0.84*	–	–	–	–	–	–	–	–	–	–
SAT9		S 99	NCE	14	0.71*	–	–	–	–	–	–	–	–	–	–
SAT9		F 00	SS	–	–	45	0.85*	–	–	–	–	–	–	–	–
SAT9		S 01	NCE	45	0.71*	105	0.81*	11	0.69*	–	–	–	–	–	–
SAT9		S 01	SS	54	0.76*	109	0.69*	19	0.27	77	0.59*	67	0.76*	71	0.65*
SAT9		F 01	SS	104	0.84*	–	–	–	–	–	–	–	–	–	–
TerraNova															
TerraNova		S 99	NCE	35	0.61*	47	0.62*	–	–	–	–	–	–	–	–
TerraNova		S 00	SS	18	0.73*	–	–	–	–	–	–	–	–	–	–
TerraNova		S 01	NCE	17	0.29	17	0.52*	–	–	–	–	–	–	–	–
TerraNova		S 01	SS	–	–	99	0.74*	–	–	–	–	–	–	–	–
TerraNova		F 01	SS	–	–	38	0.74*	–	–	–	–	–	–	–	–
Test of Achievement Proficiency															
TAP		F 01	NCE	–	–	–	–	8	0.70	7	0.70	–	–	–	–

Table 30: Other External Validity Data—STAR Math 2.0 Correlation Coefficients (r) with External Tests Administered Prior to Spring 2002, Grades 7–12[a] (Continued)

Test	Version	Date	Score	7		8		9		10		11		12	
				n	r	n	r	n	r	n	r	n	r	n	r
Texas Assessment of Academic Skills, 2001															
TAAS	2001	S 01	SS	66	0.44*	69	0.33*	–	–	–	–	–	–	–	–
Virginia Standards of Learning															
Virginia		S 00	SS	25	0.71*	–	–	–	–	–	–	–	–	–	–

Summary							
Grade(s)	**All**	**7**	**8**	**9**	**10**	**11**	**12**
Number of students	3,066	930	1,049	479	245	222	141
Number of coefficients	66	20	19	11	7	5	4
Average validity	–	0.67	0.65	0.56	0.67	0.66	0.60
Overall average	0.64						

a. n = Sample size.
* Denotes correlation coefficients that are statistically significant at the 0.05 level.

Meta-Analysis of the STAR Math Validity Data

Meta-analysis is a set of statistical procedures that combines results from different sources or studies. When applied to a set of correlation coefficients that estimate test validity, meta-analysis combines the observed correlations and sample sizes to yield estimates of overall validity, as well as standard errors and confidence intervals, both overall and within grades.

To conduct a meta-analysis of the STAR Math validity data, the 568 correlations reported in the current manual were combined and analyzed using a fixed effects model for meta-analysis. The results are displayed in Table 31. The table lists results for the correlations within each grade, as well as results with all twelve grades' data combined. For each set of results, the table lists an estimate of the true validity, a standard error, and the lower and upper limits of a 95 percent confidence interval for the validity coefficient. Using the 568 correlation coefficients, the overall estimate of the validity of STAR Math is 0.69, with a standard error of 0.001. The true validity is estimated to lie within the range of 0.69 to 0.70, with a 95 percent confidence level. The probability of observing the 568 correlations reported in Tables 25–28, if the true validity were zero, is virtually zero. Because the 568 correlations were obtained with widely different tests, and among students from twelve different grades, these results provide support for the validity of STAR Math as a measure of math skills.

Table 31: Results of the Meta-Analysis of STAR Math Correlations with Other Tests

Grade	Effect Size		95% Confidence Interval	
	Validity Estimate	Standard Error	Lower Limit	Upper Limit
1	0.55	0.01	0.53	0.57
2	0.63	0.01	0.62	0.64
3	0.68	0.00	0.67	0.68
4	0.69	0.00	0.68	0.70
5	0.71	0.00	0.70	0.72
6	0.72	0.00	0.71	0.73
7	0.73	0.01	0.72	0.74
8	0.73	0.01	0.72	0.75
9	0.74	0.02	0.70	0.77
10	0.76	0.02	0.73	0.80
11	0.78	0.02	0.73	0.82
12	0.76	0.03	0.70	0.82
All Grades	0.69	0.00	0.69	0.70

Relationship of STAR Math Scores to Teacher Ratings

In order to have a common measure of each student's math skills independent of STAR Math, Renaissance Learning constructed two 12-item checklists for teachers to use during the norming study.

On this worksheet, teachers were asked to rate each student's ability to complete a wide range of tasks related to developing math skills. The intent of this checklist was to provide teachers with a single, brief instrument they could use to rate any student.

For simplicity, two rating forms were developed: one for grades 1–5, and another for grades 6–12. This section presents the skills rating instrument itself, its psychometric properties as observed in the norming study, and the relationship between student skills ratings on the instrument and their Scaled Scores on STAR Math.

The Rating Instruments

To gather ratings of math skills from teachers, these instruments were intended to specify a sequence of skills that the teacher could quickly assess for each student and were ordered such that a student who could correctly perform the n^{th} skill in the list could almost certainly perform all of the preceding skills correctly as well. Such a list, even though quite short, provided a reliable method for sorting students from first through twelfth grade into an ordered set of math skill categories.

To construct the two ratings instruments, nineteen skill-related items were written, ranked from easiest to hardest, and assembled into two rating instruments. The first twelve items—the twelve easiest skills—formed the rating instrument used for grades 1 to 5. The eighth through nineteenth items—the twelve hardest skills—made up the instrument used for grades 6–12.

Each teacher was asked to dichotomously rate his or her students participating in the STAR Math norming study on each skill using the rating form appropriate to the student's grade. To assist with this process, the norming study software incorporated a feature enabling it to print a ratings worksheet for each participating grade. The printed ratings worksheet consisted of a checklist of the twelve skill-related performance tasks, pre-printed with the names of the participating students. To complete the instrument, the teacher had to simply mark, for each student, any task he or she believed the student could perform. The items forming both rating forms are shown on the following two pages.

Grade 1–5 Math Skills Rating Worksheet

STAR Math Norming for Grades 1–5

Sorted by: Student Name **School Name:** ______________________________

Primary Contact: ____________________________

In the table below, please identify which of the following tasks each of your students can probably do correctly.

1. Identify the longest pencil among 3 pencils of different lengths.
2. Add 2 to 4.
3. State how many cents a dime is worth.
4. Determine the number that shows "ones" in 162.
5. Subtract 7 from 35.
6. Determine the number that follows in the sequence 2, 6, 10, 14, ____.
7. Divide 18 by 3.
8. Write 78,318 in expanded form.
9. Read aloud the word name for 0.914.
10. Solve the problem 4/9 + 8/9.
11. Translate the statement "36 divided by a number is 12" into an equation.
12. Divide 11,540 by 577.

Renaissance Learning, Inc. and its subsidiaries maintain high standards of confidentiality with all data acquired for research and development purposes. Renaissance Learning assures you that all school and student data derived from these activities will only be used for research and development purposes that are intended to validate and/or improve design specifications for general product release into the education market. Individual teacher and student names, grades, and ages will be kept strictly confidential; access to this data will be limited to personnel with relevant research and development responsibilities.

		Mark an "X" for the tasks that each student probably <u>can</u> do correctly and an "O" for the tasks that each student probably <u>cannot</u> do correctly:												
Student No.	**Student Name**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**Not Rated**
1	Bartles, Amanda													
2	Bowers, Erica													
3	Driggon, Haley													
4	Edmond, Mason													
5	Edwards, Robert													
6	Halstead, Matthew													
7	Jackson, Wesley													
8	Kendricks, Marcy													
9	Lyons, Freda													

Grade 6–12 Math Skills Rating Worksheet

STAR Math Norming for Grades 6–12

Sorted by: Student Name

School Name: ______________________________

Primary Contact: ____________________________

In the table below, please identify which of the following tasks each of your students can probably do correctly.

1. Write 78,318 in expanded form.
2. Read aloud the word name for 0.914.
3. Solve the problem 4/9 + 8/9.
4. Translate the statement “36 divided by a number is 12” into an equation.
5. Divide 11,540 by 577.
6. Solve a word problem requiring the calculation of proportions.
7. Solve the problem “14 is 50% of what number?”
8. Solve a word problem requiring the calculation of 80% of 112.
9. Simplify the expression (x + 1)(x + 4)
10. Solve the equation $x^2 = 16x$.
11. Calculate vertical and supplementary angles.
12. Determine 6^{-2}.

Renaissance Learning, Inc. and its subsidiaries maintain high standards of confidentiality with all data acquired for research and development purposes. Renaissance Learning assures you that all school and student data derived from these activities will only be used for research and development purposes that are intended to validate and/or improve design specifications for general product release into the education market. Individual teacher and student names, grades, and ages will be kept strictly confidential; access to this data will be limited to personnel with relevant research and development responsibilities.

		Mark an “X” for the tasks that each student probably can do correctly and an “O” for the tasks that each student probably cannot do correctly:												
Student No.	**Student Name**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**Not Rated**
1	Bailey, Amanda													
2	Blake, Erica													
3	Duey, Haley													
4	Eaton, Mason													
5	Erlings, Robert													
6	Gable, Matthew													
7	James, Wesley													
8	Koore, Marcy													
9	Lipton, Freda													

Psychometric Properties of the Skills Ratings

Teachers completed skills ratings for 17,326 of the 29,185 students in the US norms group. The skills rating items were calibrated on an IRT scale using the Rasch model, with item parameters from both levels placed on a common scale. This allowed the skills ratings for students at both levels to be assigned a score on the same Rasch metric.

The resulting Rasch scores ranged from –14.47 to 11.1. The lower value corresponds to students in grades 1 to 5 rated as possessing none of the math skills, and the higher value corresponds to students in grades 6–12 rated as possessing all of them. Table 32 lists data about the psychometric properties of the rating scale, overall and by grade, including the correlations between skills ratings and STAR Math Scaled Scores. The internal consistency reliability of the rating scale was estimated as 0.93, using Cronbach's alpha.

Table 32: Psychometric Characteristics of the Skills Rating Scale and its Relationship to Scaled Scores, by Grade

Grade	N	Skills Rating		STAR Math Scaled Score		Correlation of Skills Ratings and Scaled Scores[a]
		Mean	S.D.	Mean	S.D.	
1	1,916	–6.60	2.95	385	89	0.40*
2	2,043	–3.67	2.41	503	84	0.47*
3	1,817	0.04	3.06	589	87	0.52*
4	1,820	1.26	2.83	651	90	0.58*
5	2,072	2.97	2.84	713	97	0.50*
6	1,637	5.50	2.07	763	100	0.44*
7	1,465	5.57	2.18	785	109	0.50*
8	1,639	6.96	2.50	811	117	0.54*
9	1,036	6.88	2.87	798	110	0.52*
10	688	8.78	2.38	824	119	0.38*
11	737	9.81	2.30	847	123	0.39*
12	456	10.03	2.05	876	127	0.42*
Overall	17,326	2.42	5.60	672	177	0.85*

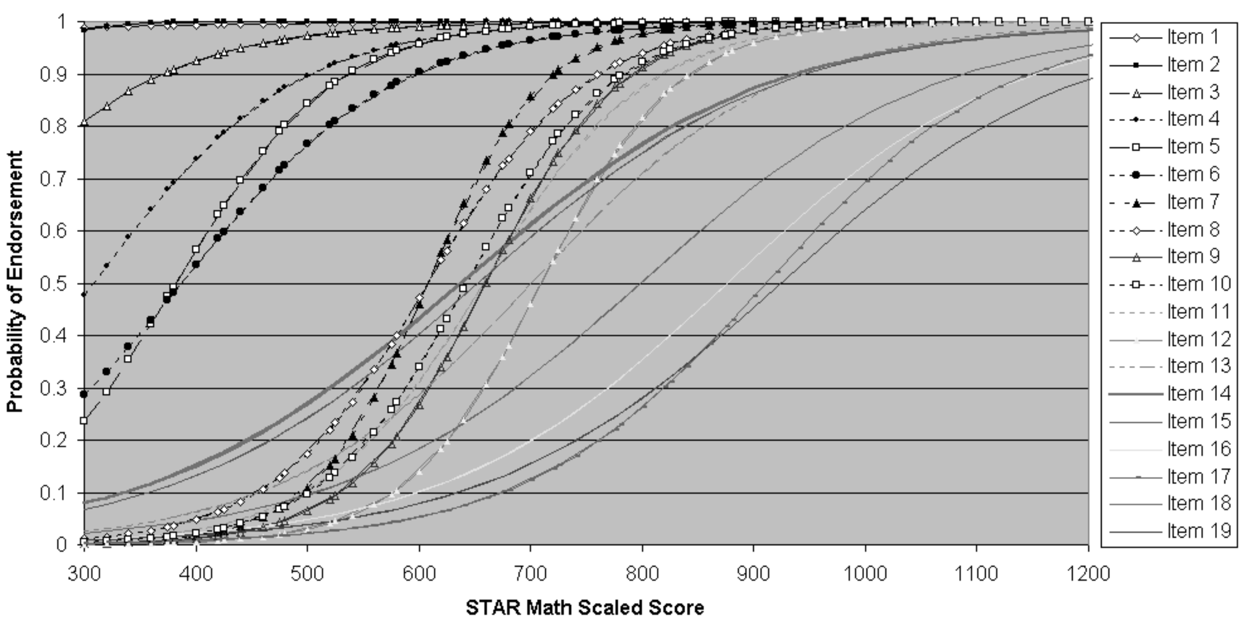
a. Asterisks denote correlation coefficients that are statistically significant at the 0.05 level.

Relationship of STAR Math Scaled Scores to Math Skills Ratings

As the data in Table 32 on page 80 show, the mean rating Scaled Scores increased directly with grade, from 6.6 at grade 1 to 10.03 at grade 12. The correlation between the skills ratings and STAR Math Scaled Scores was significant at every grade level. The overall correlation was 0.85, indicating a substantial degree of relationship between the computer-adaptive STAR Math test and teachers' ratings of their students' math skills.

Figure 4 displays the relationships of each of the nineteen rating scale items to STAR Math Scaled Scores. These relationships were obtained by fitting mathematical models to the response data for each of the rating items. Each of the curves in the figure is a graphical depiction of the respective model. As the curves show, the proportion of students rated as possessing each of the 19 rated skills increases with the STAR Math Scaled Score.

Figure 4: The Relationship of Teachers' Ratings of Student Math Skills to STAR Math Scaled Scores



The relative positions of the curves provide one indication of the relative difficulty of the 19 rated skills. The rating items' Rasch difficulty parameters, displayed in Table 33 on the next page, provide a somewhat different indication; the skills rating items are listed in the table from easiest to most difficult, by Rasch difficulty. The first column of Table 33 indicates the relative difficulty of the nineteen rating items, where relative difficulty 1 is the easiest and 19 is most difficult. The second and third columns list the item numbers and text of the skills rating items. The fourth column lists the Rasch difficulty scale value for each item. The fifth column lists the correlations between students' ratings and their STAR Math Scaled Scores.

Table 33: The Nineteen Rating Scale Items Listed in Order of Difficulty with Rasch Difficulty Parameters

Relative Difficulty	Item	Rating Scale Item	Rasch Difficulty	Correlation with Scaled Score[a]
Easiest	1	Identify the longest pencil among 3 pencils of different lengths.	–14.58	0.06*
	2	Add 2 to 4.	–14.30	0.09*
	3	State how many cents a dime is worth.	–10.28	0.26*
	4	Determine the number that shows “ones” in 162.	–7.26	0.43*
	5	Subtract 7 from 35.	–6.12	0.55*
	6	Determine the number that follows in the sequence 2, 6, 10, 14, ____.	–5.42	0.49*
	7	Divide 18 by 3.	–1.85	0.71*
	8	Write 78,318 in expanded form.	1.22	0.67*
	10	Solve the problem 4/9 + 8/9.	2.09	0.70*
	9	Read aloud the word name for 0.914.	2.51	0.70*
	11	Translate the statement “36 divided by a number is 12” into an equation.	2.59	0.67*
	12	Divide 11,540 by 577.	3.89	0.68*
	14	Solve the problem “14 is 50% of what number?”	4.54	0.40*
	15	Solve a word problem requiring the calculation of 80% of 112.	4.75	0.34*
	13	Solve a word problem requiring the calculation of proportions.	5.12	0.35*
	18	Calculate vertical and supplementary angles.	6.85	0.35*
	16	Simplify the expression (x + 1)(x + 4).	8.10	0.37*
	19	Determine 6^{-2}.	9.03	0.36*
Most Difficult	17	Solve the equation $x^2 = 16x$.	9.12	0.33*

a. Asterisks denote correlation coefficients that are statistically significant at the 0.05 level.

Notice that the first two rating scale items (“Identify the longest pencil among 3 pencils of different lengths” and “Add 2 to 4”) had extremely low Rasch difficulty indices, and correlations with Scaled Scores that were near zero. As can be seen in Figure 4 on page 81, these items were endorsed for nearly 100% of the students, regardless of their STAR Math Scaled Scores.

As a result, they did not discriminate among students with high and low levels of developed math ability, as measured by the STAR Math test.

Although teachers endorsed items 3–6 somewhat less often than items 1 and 2, they still considered these math tasks relatively easy for their students to complete. The correlations with STAR Math Scaled Scores for items 3–6 were higher than those for the first two items, but still only moderate. This may have occurred because the skills associated with items 3–6 are almost completely mastered (defined as 80% proficiency) by a student obtaining a STAR Math Scaled Score of 500.

Teachers' responses to items 7–12 suggest that their corresponding math tasks are considerably more difficult for their students to complete. This is reflected both in their Rasch difficulty parameters in Table 33 on page 82 and in Figure 4 on page 81. The figure suggests that mastery of these skills occurs between 700 and 800 on the STAR Math Score Scale. The slopes of the curves for these are all steep relative to other skills items, suggesting that these skills develop rapidly, compared to the others. The correlations between these items and Scaled Scores support this hypothesis, as items 7–12 show the highest correlations with STAR Math Scaled Scores.

Items 13–19 measure the most difficult of the skills. This is indicated by their Rasch difficulty parameters in Table 33 and is also confirmed by the locations at which 80% mastery occurs, illustrated in Figure 4, which suggests that these skills develop much later than all others. In fact, all students may not master these skills. Moreover, all of these items have only moderate correlations with STAR Math Scaled Scores, suggesting that growth of these skills is relatively gradual.

Linking STAR and State Assessments: Comparing Student- and School-Level Data

With an increasingly large emphasis on end-of-the-year summative state tests, many educators seek out informative and efficient means of gauging student performance on state standards—especially those hoping to make instructional decisions before the year-end assessment date.

For many teachers, this is an informal process in which classroom assessments are used to monitor student performance on state standards. While this may be helpful, such assessments may be technically inadequate when compared to more standardized measures of student performance. Recently the assessment scale associated with STAR Math has been linked to the scales used for summative mathematics tests in approximately 30 states, a number that is expected to increase in the near future. Linking STAR Math assessments to state tests allows educators to reliably predict student performance on their state assessment using STAR Math scores. More specifically, it places teachers in a position to identify

- which students are on track to succeed on the year-end summative state test, and
- which students might need additional assistance to reach proficiency.

Educators using STAR Math Enterprise assessments can access STAR Performance Reports that allow access to students' Pathway to Proficiency. These reports indicate whether individual students or groups of students (by class, grade, or demographic characteristics) are likely to be on track to meet a particular state's criteria for mathematics proficiency. In other words, these reports allow instructors to evaluate student progress toward proficiency and make data-based instructional decisions well in advance of the annual state tests. Additional reports automatically generated by STAR Math help educators screen for later difficulties and progress monitor students' responsiveness to interventions.

An overview of two methodologies used for linking STAR Math to state assessments is provided in the following sections.

Methodology Comparison

Recently, Renaissance Learning has developed linkages between STAR Math Scaled Scores and scores on the accountability tests of a number of states. Depending on the kind of data available for such linking, these linkages have been accomplished using one of two different methods. One method used student-level data, where both STAR and state test scores were available for the same students. The other method used school-level data; this method was applied when approximately 100% of students in a school had taken STAR Math, but individual students' state test scores were not available.

Student-Level Data

Using individual data to link scores between distinct assessments is commonly used when student-level data are readily available for both assessments. In this case, the distribution of standardized scores on one test (e.g. percentile ranks) may be compared to the distribution of standardized scores on another test in an effort to establish concordance. Recently, the release of individual state test data for linking purposes allowed for the comparison of STAR assessments to state test scores for several states. STAR test comparison scores were obtained within an eight-week window around the median state test date (+/–4 weeks).

Typically, states classify students into one of three, four, or five performance levels on the basis of cut scores (e.g. Below Basic, Basic, Proficient, or Advanced). After each testing period, a distribution of students falling into each of these categories will always exist (e.g. 30% in Basic, 25% in Proficient, etc.). Because STAR data were available for the same students who completed the state test, the distributions could be linked via equipercentile linking analysis (see Kolen & Brennan, 2004) to scores on the state test. This process creates tables of approximately equivalent scores on each assessment, allowing for the lookup of STAR scale scores that correspond to the cut scores for different performance levels on the state test. For example, if 20% of students were "Below Basic" on the state test, the lowest STAR cut score would be set at a score that partitioned only the lowest 20% of scores.

School-Level Data

While using student-level data is still common, obstacles associated with individual data often lead to a difficult and time-consuming process of obtaining and analyzing data. In light of the time-sensitive needs of schools, obtaining student-level data is not always an option. As an alternative, school-level data may be used in a similar manner. These data are publicly available, thus making the linking process more efficient.

School-level data were analyzed for some of the states included in the student-level linking analysis. In an effort to increase sample size, the school-level data presented here represent "projected" Scaled Scores. Each STAR score was projected to the mid-point of the state test administrations window using decile-based growth norms. The growth norms are both grade- and subject-specific and are based on the growth patterns of more than one million students using STAR assessments over a three-year period. Again, the linking process used for school-level data is very similar to the previously described process—the distribution of state test scores is compared to projected STAR scores and using the observed distribution of state-test scores, equivalent cut scores are created for the STAR assessments (the key difference being that these comparisons are made at the group level).

Accuracy Comparisons

Accuracy comparisons between student- and school-level data are particularly important given the marked resource differences between the two methods. These comparisons are presented for three states[4] in Tables 34–36. With few exceptions, results of linking using school-level data were nearly identical to student-level data on measures of specificity, sensitivity, and overall accuracy. McLaughlin and Bandeira de Mello (2002) employed similar methods in their comparison of NAEP scores and state assessment results, and this method has been used several times since then (McLaughlin & Bandeira de Mello, 2003; Bandeira de Mello, Blankenship, & McLaughlin, 2009; Bandeira et al., 2008).

4. Data were available for Arkansas, Florida, Idaho, Kansas, Kentucky, Mississippi, North Carolina, South Dakota, and Wisconsin; however, only North Carolina, Mississippi, and Kentucky are included in the current analysis.

In a similar comparison study using group-level data, Cronin et al. (2007) observed cut score estimates comparable to those requiring student-level data.

Table 34: Number of Students Included in Student-Level and School-Level Linking Analyses by State, Grade, and Subject

State	Grade	Math	
		Student	School
NC	3	1,100	524
	4	751	890
	5	482	551
	6	202	515
	7	216	67
	8	39	372
MS	3	1,786	4,309
	4	1,757	4,584
	5	1,531	5,294
	6	1,180	5,190
	7	721	3,390
	8	549	1,896
KY	3	3,777	935
	4	3,155	1,797
	5	2,228	1,430
	6	1,785	1,497
	7	788	984
	8	362	1,036

Table 35: Comparison of School Level and Student Level Classification Diagnostics for Mathematics

State	Grade	Sensitivity[a]		Specificity[b]		False + Rate[c]		False – Rate[d]		Overall Rate	
		Student	School	Student	School	Student	School	Student	School	Student	School
NC	3	92%	81%	53%	73%	47%	27%	8%	19%	80%	78%
	4	90%	78%	52%	73%	48%	27%	10%	22%	80%	78%
	5	83%	83%	62%	57%	38%	43%	17%	17%	75%	74%
	6	94%	87%	42%	65%	58%	35%	6%	13%	74%	83%
	7	91%	88%	61%	69%	39%	31%	9%	12%	81%	84%
	8	89%	77%	58%	76%	42%	24%	11%	23%	77%	77%
MS	3	78%	70%	77%	83%	23%	17%	22%	30%	77%	76%
	4	73%	73%	81%	81%	19%	19%	27%	27%	77%	77%
	5	71%	68%	83%	84%	17%	16%	29%	32%	77%	76%
	6	71%	66%	81%	85%	19%	15%	29%	34%	76%	76%
	7	83%	84%	82%	81%	18%	19%	17%	16%	83%	83%
	8	56%	66%	89%	83%	11%	17%	44%	34%	76%	76%
KY	3	95%	92%	45%	54%	55%	46%	5%	8%	83%	83%
	4	92%	87%	47%	60%	53%	40%	8%	13%	80%	80%
	5	90%	90%	51%	50%	49%	50%	10%	10%	77%	77%
	6	82%	80%	64%	68%	36%	32%	18%	20%	75%	75%
	7	72%	68%	81%	85%	19%	15%	28%	32%	76%	76%
	8	59%	66%	89%	85%	11%	15%	41%	34%	74%	76%

a. Sensitivity refers to the proportion of correct positive predictions.
b. Specificity refers to the proportion of negatives that are correctly identified (e.g. student will not meet a particular cut score).
c. False + rate refers to the proportion of students incorrectly identified as “at-risk.”
d. False – rate refers to the proportion of students incorrectly identified as *not* “at-risk.”

Table 36: Comparison of Differences Between Achieved and Forecasted Performance Levels in Math (Forecast % – Achieved %)

State	Grade	Student	School	Student	School	Student	School	Student	School
NC		Level I		Level II		Level III		Level IV	
	3	–2.6%	–1.6%	–2.8%	0.8%	15.6%	2.1%	–10.2%	–1.3%
	4	–4.0%	–0.4%	–2.5%	1.2%	14.7%	1.5%	–8.2%	–2.3%
	5	–2.7%	–0.9%	1.6%	–3.9%	10.0%	11.6%	–8.9%	–6.7%
	6	–7.3%	–5.3%	–8.2%	–4.5%	18.6%	7.1%	–3.1%	2.7%
	7	–1.3%	–0.6%	–5.0%	–1.1%	15.1%	1.1%	–8.8%	0.6%
	8	–4.2%	–4.4%	–5.6%	–2.9%	2.5%	–1.2%	7.4%	8.6%
MS		Minimal		Basic		Proficient		Advanced	
	3	2.7%	10.1%	0.0%	0.2%	1.1%	–15.0%	–3.9%	4.6%
	4	1.5%	9.9%	4.4%	–3.4%	–3.7%	–10.7%	–2.1%	4.2%
	5	0.8%	9.4%	5.3%	–1.0%	–3.5%	–11.3%	–2.7%	2.8%
	6	4.7%	12.6%	–0.8%	–4.3%	–1.8%	–11.6%	–2.1%	3.3%
	7	0.7%	2.8%	–0.5%	–3.7%	0.0%	–1.8%	–0.2%	2.8%
	8	5.8%	7.0%	4.6%	–4.4%	–9.9%	–4.1%	–0.5%	1.5%
KY		Novice		Apprentice		Proficient		Distinguished	
	3	–3.2%	–2.0%	–4.8%	–2.6%	12.1%	3.3%	–4.0%	1.4%
	4	–4.1%	–2.7%	–3.9%	1.0%	5.6%	1.6%	2.4%	0.1%
	5	–3.7%	–0.2%	–5.4%	–9.7%	11.4%	8.4%	–2.3%	1.6%
	6	–3.9%	–0.4%	0.1%	–0.5%	5.8%	0.5%	–2.1%	0.2%
	7	–1.9%	7.1%	10.5%	3.6%	1.2%	–3.0%	–9.6%	–7.5%
	8	1.5%	4.3%	13.8%	4.9%	–5.0%	–1.9%	–10.2%	–7.3%

Norming

Versions of STAR Math released between 2002 and 2011, including STAR Math Enterprise, use the STAR Math version 2 Scaled Score norms developed in 2002. In 2012, updated test score norms were computed for the STAR Math Service version, for introduction at the beginning of the 2012–13 school year. This chapter describes the 2012 norming of the STAR Math Service version.

In addition to Scaled Score norms, Renaissance Learning has developed growth norms for STAR Math. The section on growth norms in this chapter describes the development and use of the growth norms, which have been in use since 2008. Growth norms are very different from test score norms, having different meaning and different uses. Users interested in growth norms should familiarize themselves with the differences, which are made clear in the growth norms section (see page 97).

Sample Characteristics

Students' STAR Math data in the Renaissance Learning Hosted Learning Environment ranging from fall 2008 to spring 2011 were used for the 2012 STAR Math norming study. The 2012 STAR Norming Sample included students from 48 US states and the District of Columbia. The US states not represented in the 2012 norming sample were Rhode Island and Vermont. School and district demographic data were obtained from Market Data Retrieval (MDR), National Center for Education Statistics (NCES), and the US Bureau of Census. Students' demographic data when recorded included Gender, Race/Ethnicity, Bilingual Status, Free Lunch, Reduced Lunch, Learning Disability, Physical Disability, English Language Learner, Gifted and Talented, Limited English Proficient, Title 1, and Special Education.

To obtain a representative sample of the US school population, a multi-stage stratified random sampling process was used. The stratification variables are described below. The first sampling stage selected representative samples from different geographic regions (East, Midwest, West) and metropolitan classification codes (rural, suburban, urban).The second sampling stage selected representative samples from different school sizes and socioeconomic status classifications. Socioeconomic status included four classification levels for the percent of students in the school that qualified for free and reduced student lunch. The third sampling stage selected representative samples from grades 1–10 and ten deciles (deciles 1–10 of STAR Math scores) within each grade. From the norming sample completed in the first three stages described above, the fourth and final sampling stage selected equal sample sizes from the last three years of STAR Math data (fall 2008–spring 2009, fall 2009–spring 2010, and fall 2010–spring 2011). The fourth and final sampling stage merely assured representative sampling from the last three years of STAR Math data.

The key stratification variables were:

Geographic Region. Using the categories established by the National Center for Education Statistics (NCES), students were grouped into three geographic regions: East (including Northeast and Southeast), Midwest, and West.

East

Alabama, Arkansas, Connecticut, Delaware, District of Columbia, Florida, Georgia, Kentucky, Louisiana, Maine, Maryland, Massachusetts, Mississippi, New Jersey, New Hampshire, New York, North Carolina, Pennsylvania, Rhode Island, South Carolina, Tennessee, Vermont, Virginia, and West Virginia.

Midwest

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

West

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

School Metropolitan Classification. Using the categories from Market Data Retrieval (MDR), schools were classified as rural (non-metropolitan), suburban, and urban schools. Rural schools are classified as schools with rural and non-metropolitan postal ZIP codes that do not fall within the boundaries of a Metropolitan Area (MA). Suburban schools have postal ZIP codes that fall within the geographical confines of an MA, but fall outside the central cities. Urban schools have postal ZIP codes that include the central city that gives its name to the MA.

School Size. Based on total school enrollment, schools were classified into one of three school size groups: small schools had under 500 students enrolled, medium schools had between 500–999 students enrolled, and large schools had 1,000 or more students enrolled.

Socioeconomic Status as Indexed by the Percent of School Students with Free and Reduced Lunch. Schools were classified into one of four classifications based on the percentage of students in the school who had free or reduced student lunch. The classifications were coded as follows:

1 High Socioeconomic Status (0%–24%)

2 Above Median Socioeconomic Status (25%–49%)

3 Below Median Socioeconomic Status (50%–74%)

4 Low Socioeconomic Status (75%–100%)

No students were sampled from the school classifications that did not report the percent of school students with Free and Reduced Lunch. The implication of this factor for the norming cannot be determined. The norming sample also included many private and parochial schools as described below.

Grade. The STAR Math 2012 norming sample comprised students from grades 1–10. There was insufficient data for sampling students and computing norms for Kindergarten and grades 11 and 12.

Deciles. Students' STAR Math scale scores were grouped into 10 deciles from the fall 2008–spring 2011 data and then students were randomly sampled from each of the ten deciles classifications within each grade level.

School Year. Data were selected from fall 2008–spring 2011, with equal samples drawn from each school year.

Tables 37 to 41 summarize some key variables from the fall 2008 to spring 2011 norming sample.

Table 37: Sample Characteristics, STAR Math Norming Study—Fall 2008–Spring 2011 (N = 450,007 Students)

	Students	
	National %	Sample %
Geographic Region		
East	53.92%	51.75%
Midwest	21.49%	21.33%
West	24.59%	26.92%
District Socioeconomic Status (Percentage of Free/Reduced Lunch)		
High (0%–24%)	25.3%	23.60%
Above Median (25%–49%)	26.3%	24.47%
Below Median (50%–74%)	24.8%	25.21%
Low (75%–100%)	22.1%	26.73%
School Size		
1–599 Students	45.30%	46.38%
600–999 Students	42.30%	58.63%
1,000+ Students	12.40%	4.98%

Table 38: School Locations, STAR Math Norming Study—Spring 2012 (N = 450,007 Students)

	Students	
	National %	Sample %
Rural	37.25%	34.01%
Suburban	36.10%	33.24%
Urban	26.65%	32.76%
Total	100.00%	100.00%

Table 39: Gender and Ethnic Group Participation, STAR Math Norming Study—Spring 2012 (N = 450,007 Students)

		Students	
		National %	Sample %
Ethnic Group	Asian/Pacific Islander	4.3%	2.72%
	Black	14.1%	19.51%
	Hispanic	21.8%	10.02%
	Native American	0.9%	4.11%
	White	56.1%	39.36%
	Other	3.0%	0.63%
	Unrecorded	N/A	69.03%[a]
Gender	Female	48.95%	38.18%
	Male	51.05%	39.14%
	Unrecorded	N/A	25.68%

a. The data for ethnic group participation should not be considered representative of the US population since there was only a 30% response rate for ethnic group recording.

Table 40: Type of School

	National %	Sample %
Public & Charter	80.3%	90.0%
Private	13.7%	2.3%
Catholic	6.1%	4.1%
Other[a]	–	3.7%
All Types	100%	100%

a. Other schools in the sample included state-operated schools (3.0%), county-operated schools (0.13%), colleges (0.01%), regional centers (0.0%, 10 regional center schools) and Bureau of Indian Affairs schools (0.47%).

The STAR Math 2012 norming sample included 89.96% public schools, 4.14% Catholic schools, 3.00% state-operated schools, 2.29% private schools, 0.47% Bureau of Indian Affairs schools, 0.13% county-operated schools, 0.01% district schools, 0.01% schools affiliated with colleges, and ten schools (0.00%) associated with regional centers.

Table 41: District/School Poverty Level Code

District Poverty Level Code	National Districts %	National Schools %	Sample %
A 0%–5.9%	13.2%	10.8%	2.2%
B 6%–15.9%	43.6%	41.1%	33.4%
C 16%–30.9%	37.2%	42.5%	50.3%
D 31% or More	6.0%	5.7%	11.9%
E Unclassified	–	–	2.1%
Total	100.0%	100.0%	100.0%

The STAR Math 2012 norming sample included 76 bilingual students, 6,531 students who qualified for free lunch, 417 students with learning disabilities, 59 students with physical disabilities, 1,579 students who were English Language Learners (ELL), 1,946 students who were gifted and talented (G&T), 2,740 Title I students, and 3,117 Special Education students.

Data Analysis

After selecting a stratified random sample of US students from grades 1–10, sample characteristics were summarized to determine the degree of correspondence to the national population. These sample summaries are shown in Tables 37 and 41. Unweighted scores were used for compiling the norms due to the similarity of the sample proportions to the national population proportions based on the characteristics of geographic region, socioeconomic status, school size and school location. Due to the high proportion of missing data for gender and ethnic group participation, the norming sample proportions should not be considered as representative of the national population.

Both fall and spring scores were used in the norming study. Table 42 shows the fall 2008 to fall 2011 Scale Score summary statistics by grade whereas Table 43 shows the spring 2008 to spring 2011 Scale Score summary statistics, also by grade.

Table 42: Comparison of Scaled Scores, STAR Math Norming Study—Fall 2008–Fall 2011 (N = 425,007 Students)

Grade	Sample Size	Scaled Score Means	Scaled Score Standard Deviations	Scaled Score Medians	Minimum Scaled Score	Maximum Scaled Score
1	20,240	267	93	263	1	813
2	53,422	408	87	414	1	811
3	91,485	495	86	500	1	937
4	80,970	579	92	585	82	1,007

Table 42: Comparison of Scaled Scores, STAR Math Norming Study—Fall 2008–Fall 2011 (N = 425,007 Students) (Continued)

Grade	Sample Size	Scaled Score Means	Scaled Score Standard Deviations	Scaled Score Medians	Minimum Scaled Score	Maximum Scaled Score
5	69,478	645	98	650	1	1,064
6	47,215	711	103	718	68	1,112
7	30,360	747	110	757	125	1,187
8	21,450	777	118	790	123	1,318
9[a]	6,105	790	117	802	180	1,215
10[a]	4,462	793	123	806	152	1,337

a. Grades 9 and 10 had substantially lower sample sizes.

Table 43: Comparison of Scaled Scores, STAR Math Norming Study—Spring 2008–Spring 2011 (N = 425,007 Students)

Grade	Sample Size	Scaled Score Means	Scaled Score Standard Deviations	Scaled Score Medians	Minimum Scaled Score	Maximum Scaled Score
1	20,240	406	91	406	1	813
2	53,422	514	86	513	1	980
3	91,485	597	93	605	1	991
4	80,790	656	97	663	1	1,078
5	69,478	710	100	717	72	1,192
6	47,215	763	106	769	122	1,279
7	30,360	785	114	794	100	1,379
8	21,450	813	123	819	90	1,374
9[a]	6,105	819	118	822	58	1,256
10[a]	4,462	823	127	828	90	1,289

a. Grades 9 and 10 had substantially lower sample sizes.

The sample sizes per grade for Tables 42 and 43 are identical because students were selected for the norming sample if there were matched fall and spring scores from the same students.

The norm-referenced scores are determined from both the fall and spring testing periods used for the norming. The date range for the fall scores was August 1 to October 15 of the school year, and the spring scores were obtained between April 15 and the end of school year. For the STAR Math 2012 norms, September was selected as the testing month for fall scores, and June was selected for the spring scores. Scores were linearly interpolated between fall (September) and spring (June) assuming equal growth for each of the ten school months (September–June) and no expected growth for the summer months of July and August. Summer norms were not computed.

Grade Equivalent (GE) scores for each grade and each month of the school year were computed using the median STAR Math scaled scores for each grade and month. The Scaled Score to Grade Equivalent conversion table is presented in Table 57 on page 128. The Scaled Score to Percentile Rank conversion tables for the empirical norming period are presented in Table 58 on page 129. This norming approach allows STAR Math to provide normative information that is most relevant, regardless of the specific time period in which schools administer the STAR Math test to students.

Additional Information Regarding the Norming Sample

Table 44 shows the frequency and percent of test records selected from each of the last three school years. This table shows that 141,669 cases were selected from the sample for each school year.

Table 44: Frequency and Percent of STAR Mathematics Records by School Year Included in the STAR Math 2012 Spring Norm Sample (N = 425,007 Students)

School Year	Frequency	Percent
2008–2009	141,669	33.33%
2009–2010	141,669	33.33%
2010–2011	141,669	33.33%

Table 45 displays the frequency and percent for School Enrollment Size Code for the norms sample. Table 45 shows classifications for seven school enrollment size codes. These classifications are from Market Data Retrieval. In many Market Data Retrieval reports the seven classifications are reduced to three school-size classifications as described above.

Table 45: Frequency and Percent for School Enrollment Size Code STAR Math—Spring 2012 (N = 425,007 Students)

School Enroll Code	Frequency	Percent
A 1–99 Students	2,970	0.70%
B 100–199 Students	18,246	4.29%
C 200–299 Students	38,235	9.00%
D 300–499 Students	137,670	32.40%
E 500–999 Students	206,667	48.63%
F 1,000–2,499 Students	19,927	4.69%
G 2,500 or More Students	1,225	0.29%
Frequency Missing	67	0.02%

Table 46 shows the frequency and percent for the District Enrollment Size Code for the norms sample. This table shows the school district enrollment classification according to the seven Market Data Retrieval classifications for district enrollment of students.

Table 46: Frequency and Percent for District Enrollment Size Code STAR Math—Spring 2012 (N = 425,007 Students)

District Enrollment	Frequency	Percent
A 1–599 Students	19,369	5.07%
B 600–1,199 Students	26,189	6.85%
C 1,200–2,499 Students	53,010	13.86%
D 2,500–4,999 Students	65,001	17.00%
E 5,000–9,999 Students	73,388	19.19%
F 10,000–24,999 Students	75,093	19.64%
G 25,000 or More Students	70,328	18.39%
Frequency Missing	42,629	10.03%

Table 47 indicates the School Level and Type.

Table 47: Frequency and Percent of School Level and Type, STAR Norming Study—Spring 2012 (N = 425,007 Students)

School Type	Frequency	Percent
A Adult School	1	0.00%
C Combined School	17,612	4.14%
E Elementary School	334,156	78.63%
G College Related	28	0.01%
J Junior High School	8,511	2.00%
M Middle School	50,262	11.83%
P Special School	1,707	0.40%
S Senior High School	11,721	2.76%
V Vocational/Tech School	970	0.23%
Frequency Missing	39	0.009%

Table 48 indicates the School Administrative Classification as state, county, district, public schools, private schools, Catholic schools, colleges, Bureau of Indian Affairs and Regional Centers.

Table 48: Frequency and Percent for School Administrative Classification, STAR Norming Study—Spring 2012 (N = 425,007 Students)

School Administrative Classification		Frequency	Percent
2	State-Operated Schools	12,731	3.00%
4	County-Operated Schools	567	0.13%
5	Districts	29	0.01%
7	Public Schools	382,349	89.96%
9	Private Schools	9,724	2.29%
10	Catholic Schools	17,578	4.14%
12	Colleges	28	0.01%
13	Bureau of Indian Affairs	1,991	0.47%
14	Regional Centers	10	0.00%
Frequency Missing		0	0.00%

Growth Norms

Since it is important to measure students' growth on academic achievement over their school career, STAR assessments were constructed to provide a vertical scale that can be used to follow student growth both within an academic year and across contiguous academic years. STAR was designed specifically to allow educators to follow students' growth over time. To enhance the utility of STAR assessments for indexing growth, growth norms were developed in 2008 to allow for making norm-referenced comparisons of student absolute growth.

At present the growth norms in STAR Math are based on over 2 million student assessments (N = 2,543,319). Growth norms are a basic extension of the norming process that takes place in much of educational and psychological testing. Most tests only provide norm-referenced information with respect to a student's performance at a particular point in time, which is similar to the STAR assessment's Percentile Rank (PR) and Grade Equivalent (GE) scores that are reported with each STAR test a student takes. However, growth norms go a step beyond this traditional method by providing a reference to student growth over time within the academic year. Growth norms were developed to index growth of representative student groups. This provides a method of comparing a student's observed growth over a period of time to growth made by students of a similar grade and achievement level.

Growth norms in the STAR assessment were developed for each grade by following students across the entire academic year, ranging from August to June (depending on the specific schedule for each school). Students were tested both at the beginning and end of the school year (during the fall and spring semesters, respectively), allowing the student growth estimates to be computed across the academic year. To normalize differences in time between the initial test and the final test at the end of the school year, the measure of growth (change in score from fall to spring testing) was divided by the number of weeks between the assessment occasions to obtain an estimate of typical growth per week for all students.

Since students develop academic skills at different rates as they mature and move across the grades, they also develop and grow at different rates within each grade depending on where they score in the overall distribution of performance. Students who score in the top decile for a grade do not (and should not be expected to) grow at the same rate across the academic year as students in the middle or lower deciles, and vice versa. Therefore, it would be problematic to use growth rate expectations obtained from a group of students who were not of comparable achievement levels. Growth rates of students should be compared to students of similar academic achievement levels; otherwise, there is the potential to expect too much or too little growth from certain students.

To account for differences in student growth, both across grades and within grades during an academic year, growth norms were developed by using information about grade and level of performance to construct homogeneous student groupings for comparison. The within-grade groupings were done by partitioning students into decile groups based on their initial Percentile Rank scores within a school year; STAR Math Percentile Ranks are based on the nationally representative samples used to generate the norms, as described earlier. For example, students in different grades will have different expectations of growth, and students within the same grade can have different levels of growth depending on their performance and achievement level. Students within the same grade in the lowest decile group will have a different growth expectation from students in the other nine decile groups within that grade.

STAR growth norms were constructed by following students within each decile of each grade across the entire academic year. This provided a means to compute a distribution of growth scores for every decile group for all grades, i.e. 10 decile groups for each grade will each have their own growth norms distribution. The growth norms are thus conditional on both grade and decile level of student initial performance during the academic year.

All data was retrieved from the hosted customer database (Renaissance Place Real Time), and the growth norms are updated periodically. This allows for the norms to continually reflect the effects of changes in educational practices that can have differential effects on student learning, and also keeps the growth norms up-to-date to ensure students' observed growth is being referenced against a current student group.

Growth norms provide a norm-referenced method of computing expected growth distributions for students. This allows student growth over time to be referenced to a known normative distribution of growth. Therefore, a relative measure of student growth can be obtained by placing a student's growth over time in the context of the distribution of growth observed among students of a similar grade and initial performance level. To provide an estimate of the normal growth for students of any grade and decile level, the median growth rate from the growth rate distribution is used (see table 49). However, in principle, any observed growth for a student can be referenced against the appropriate growth distribution to evaluate where that student's growth ranks in comparison to other similar students.

Table 49: Median Weekly Growth in STAR Math Scaled Score Points, by Student's Grade and Decile Group

Grade	Decile									
	1	2	3	4	5	6	7	8	9	10
1	5.7477	4.8235	4.5957	4.3355	4.1936	3.8677	3.7421	3.3772	2.9345	1.9145
2	4.4211	3.8571	3.5135	3.3676	3.192	2.973	2.6866	2.4456	2.2757	1.6757
3	3.7806	3.5875	3.4186	3.2407	3.0873	2.9211	2.7308	2.5075	2.2453	1.6053
4	3.2595	2.9057	2.7078	2.576	2.548	2.3975	2.1927	2.0625	1.8103	1.3258
5	2.8333	2.4848	2.4343	2.2451	2.1109	2.0082	1.875	1.7364	1.4474	1.0571
6	2.3886	2.0491	1.7645	1.6504	1.5385	1.4937	1.2419	1.0413	1.0664	0.7527
7	2.008	1.4264	1.1581	1.0913	0.9655	0.7778	0.7269	0.6571	0.5558	0.1389
8	1.9679	1.3831	1.1529	1.0107	0.6734	0.6121	0.4757	0.4863	0.0802	0.0802
9	0.489	0.489	0.489	0.489	0.489	0.489	0.489	0.489	0.489	0.489
10	0.4975	0.4975	0.4975	0.4975	0.4975	0.4975	0.4975	0.4975	0.4975	0.4975
11	0.4462	0.4462	0.4462	0.4462	0.4462	0.4462	0.4462	0.4462	0.4462	0.4462
12	0.2986	0.2986	0.2986	0.2986	0.2986	0.2986	0.2986	0.2986	0.2986	0.2986

Score Definitions

Types of Test Scores

In a broad sense, STAR Math software provides two different types of test scores that measure student performance in different ways:

- *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 Numeration and Computation 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 software. Both of these scores are based on a comparison of a student's test results to the data collected during the 2002 national norming study.

Scaled Score (SS)

STAR Math software creates 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 norm-referenced scores, all STAR Math test scores are converted to a common scale, creating Scaled Scores. The STAR Math software does this in two steps. First, maximum likelihood is used to estimate each student's location on the Rasch ability scale, based on the difficulty of the items administered, and the pattern of right and wrong answers. Second, using a linear transformation to make all scores positive integers, the Rasch ability scores are converted to STAR Math Scaled Scores. STAR Math Scaled Scores range from 0 to 1400.

STAR Math Scaled Scores are expressed on the same scale used in the previous versions of STAR Math. STAR Math Scaled Scores provide a single scale for measuring the math achievement of students from first through twelfth grade. In addition, STAR Math norm-referenced scores are derived from the within-grade distributions of Scaled Scores in the norms group.

STAR Math Enterprise Scaled Scores are expressed on the same scale used for STAR Math. The scaling of STAR Math Enterprise was accomplished by two means. First, as described earlier in the section on Item and Scale Calibration, all STAR Math Enterprise items' Rasch difficulty parameters have been calibrated on the STAR Math score scale by administering them as unscored items within STAR Math tests, and calculation the logistic regression of item responses on STAR Math Rasch ability estimates. This is tantamount to pre-equating of Enterprise scores to the STAR Math scale. Second, these "pre-equated" scores from STAR Math Enterprise tests were linked to the STAR Math score scale by applying linear equating to test scores of thousands of students who took concurrent administrations of STAR Math and STAR Math Enterprise tests during the Spring of 2011. Based on those linking analyses, small adjustments are made to STAR Math Enterprise scores to place them on the STAR Math score scale before they are reported. The effect of these linking analyses is to yield STAR Math Enterprise scale scores that are distributed very similarly to STAR Math scale scores, making it possible to use STAR Math norms to interpret STAR Math Enterprise scale scores.

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 STAR Math 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.

GEs in STAR Math range from 0.0 to 12.9+. Because the GE scale expresses individual "months" in tenths, the scale does not cover the summer months. Table 50 indicates how the decimalized GE tenths correspond to the various calendar months. Since the norming of STAR Math took place during the seventh month of the 2002 school year, GEs ending in 0.7 are empirically based; in other words, they provide conversions based on actual normative medians. All other portions of the scale are formed by fitting a curve to the grade-by-grade medians and finding Scaled Scores that fit the curve. Table 57, "Scaled Score to Grade Equivalent Conversions," on page 128 contains the Scaled Score to GE conversions.

Table 50: 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.

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. Consideration of this phenomenon should be made when averaging GE scores, especially those spanning two or more grades.

Grade Equivalent Cap

For customers who are using either STAR Math or STAR Math Enterprise on the Renaissance Place Real Time hosted platform, GE scores will be capped when they exceed three grade levels above the student's actual grade placement. When a student's Scaled Score produces a GE that is greater than the start of three grades above the student's current grade, STAR Math will report that student's GE is greater than the cap grade but will not report the specific GE score. Because this cannot happen to students in tenth grade or above, the potential for a capped GE will only exist for K–9 students. When applicable, the GE cap will now appear on all STAR Math reports—even those showing test scores from tests taken prior to this update.

For example, a fourth grade student cannot receive a GE score above 7.0 at any time of the year. If their GE score is above a 7.0, the reports will show a capped GE score of ">7".

Table 51: Grade Equivalents with GE Cap

Grade Placement	Grade Equivalent	Grade Equivalent Reported As
4.6	6.9	6.9
4.6	7.0	7.0
4.6	7.1	>7

Comparing STAR Math GEs with Those from Conventional Tests

Because STAR Math adapts to the proficiency level of the student being tested, the GE scores that STAR Math provides 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 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 up to three grade levels above grade placement, depending upon his or her actual performance on the test. Throughout a STAR Math test, students are tested on items of an appropriate level of difficulty, based on their individual level of achievement.

Functional Grade Level (FGL) ENTERPRISE

Functional Grade Level (FGL) provides teachers with information as to whether a student is performing on grade level, or above or below it, with respect to mastery of grade-specific Common Core standards or skills. In this revised concept, Functional Grade Level is a classification, not a numeric score. Given a STAR Math Enterprise Scaled Score for a student who is in a specific grade, the teacher can easily classify the student as being below, on, or above grade level in terms of the student's proficiency on standards or skills that are appropriate to the student's current grade.

Functional Grade Level is a criterion-referenced classification. The criteria it employs are specific to the knowledge, skills and abilities expected at each grade level as expressed in formal sets of grade level standards or skills. For students in states that have adopted the Common Core State Standards (CCSS) (http://www.corestandards.org/assets /CCSSI_Math%20Standards.pdf), the FGL classifications are best made with reference to those standards. For students in other states, the FGL classifications might be made with reference to grade-level skills embodied in Core Progress for Math, Renaissance Learning's compilation of Math learning progressions. This section of the technical manual deals with CCSS standards only.

This concept of FGL says that a student is on grade level if the student has satisfied minimum end-of-school-year expectations for the next lower grade. For example, for a 4th-grade student to be on grade level, that student should have knowledge and skills equivalent to mastery of the 3rd-grade standards. We determine whether the student has mastered those knowledges and skills by inference, based on the student's STAR Math Enterprise Scaled Score.

While the FGL, like any test result, is an estimate, not a certainty, it provides a useful indication of the level of material on which the student should be receiving instruction in a curriculum based on CCSS. For example, if a beginning 4th-grade student receives a STAR Math FGL classification of "on grade level," this indicates that the student has likely mastered most third-grade material, and is probably prepared to learn without experiencing too many difficulties when using materials appropriate at the start of fourth grade.

In effect, the STAR Math FGL references each student's STAR Math performance to the difficulty of material appropriate for instruction. This is a valuable piece of information in planning the instructional program for individuals or groups of students.

What Does a STAR Math FGL Classification Mean?

STAR Math FGL classifies students into three broad categories of knowledge and skills, depending on their school grade and STAR Math Scaled Score. Scaled Scores are used to calculate the expected percent correct on sets of test items that measure a wide array of knowledge and skills aligned to the CCSS.

A student is considered "on grade level" if he/she appears proficient on the standards or skills of the next lower grade. For example, a 4th-grade student on the first day of the school year would be considered "on grade level" if his/her STAR Scaled Score indicated satisfactory proficiency in the knowledge and skills specified in the standards to be learned by the end of 3rd grade. A STAR Math Scaled Score lower than the proficiency threshold would classify the student as "below grade level." A score at or above the threshold for the next higher grade would earn a classification of "above grade level."

"Satisfactory proficiency" is something that Renaissance Learning cannot define for users of STAR Math Enterprise, because proficiency standards typically vary from state to state, district to district, or even school to school. However, in this section, we are defining a student as satisfactory—on grade level—if he/she has an average proficiency level of at least 70%, calculated across the previous grade's CCSS standards. Users of STAR Math Enterprise may prefer to set a different proficiency standard for "on grade level."

Table 52 displays the minimum STAR Math Enterprise Scaled Scores ("cut" scores) for classifying students as "on grade level" according the criteria applied here. Later, Tables 53 and 54 will display more detailed data, including Scaled Score ranges that define "below grade level," "on grade level," and "above grade level" classification at every grade, and at different points within the school year.

Note that these grade-level classifications are not norm-referenced like the well-known Grade Equivalent (GE) score. They are based on evidence that the student knows and can do the things that the CCSS standards expect to be taught and successfully learned at each grade level.

Table 52: Minimum Scaled Scores for CCSS On-Grade Level Classification at Each Grade, K–12

Grade	Minimum Scaled Score
K	0
1	382
2	431
3	503
4	628
5	711
6	797
7	847
8	849
9	889
10	899
11	904
12	908

Practical Impact of These Cut Scores

If these cut scores are used to classify students as "on grade level," how many students would qualify at the beginning of each school year? To answer this question, tables of

STAR Math norms for the start of the school year in grades K–12 were consulted. Table 53 displays the percentages of students who would qualify in each grade by achieving an average proficiency level of at least 70%, calculated across the CCSS standards for the next lower grade. As the table shows, the percent of students "on grade level" at the beginning of the school year may range from as little as 12% (grade 1) to as high as 49% (grade 3). It's 100% at grade K because, for kindergarten, there are no previous grade standards. These percentages may seem alarmingly low; this is a reflection of the rigor of the CCSS standards.

Table 53: The Impact of the Cut Scores in Table 52: Expected Percent of Students Qualifying as "On Grade Level" at the Beginning of the School Year by Method, Mastery Criterion, and Grade

Grade	Expected Percent On Grade Level
K	100
1	12
2	41
3	49
4	30
5	29
6	30
7	17
8	26
9	18
10	19
11	24
12	28

Cut Scores at Different Times of the School Year

Students who are on grade level are expected to improve in knowledge and skills during the course of the school year. Recognizing this, we have calculated Scaled Score ranges that represent on-pace performance during the fall trimester (the first 3 months of the school year), the winter trimester (months 4–6), and the spring trimester (months 7–9). These appear in Table 54. It lists the STAR Math Scaled Score ranges for below-, on-, and above-grade level performance for grades K–12 on the Common Core State

Standards for Math, including Scaled Score ranges for on-grade level classifications for each of the three trimesters.

Table 54: STAR Math Enterprise CCSS FGL Cut Scores Based on an Average of 70% Mastery Across Standards

			On Grade Level							
	Below Grade Level		On Pace, Fall		On Pace, Winter		On Pace, Spring		Above Grade Level	
Grade	Low	High	Low	High	Low	High	Low	High	Low	High
K	n/a	n/a	0	127	128	254	255	381	382	1400
1	0	381	382	398	399	414	415	430	431	1400
2	0	430	431	455	456	478	479	502	503	1400
3	0	502	503	544	545	586	587	627	628	1400
4	0	627	628	655	656	683	684	710	711	1400
5	0	710	711	739	740	768	769	796	797	1400
6	0	796	797	813	814	830	831	846	847	1400
7	0	846	847	847	848	848	849	849	850	1400
8	0	849	850	862	863	875	876	888	889	1400
9	0	888	889	892	893	895	896	898	899	1400
10	0	898	899	901	902	902	903	903	904	1400
11	0	903	904	905	906	906	907	907	908	1400
12	0	907	908	909	910	910	911	911	912	1400

What Does the STAR Math FGL Classification Imply for Instructional Planning?

The FGL classifies students into one of three broad categories, based on their STAR Math Scaled Scores:

1. **On grade level.** Students in this category have attained Scaled Scores that indicate they have equaled or surpassed the CCSS end-of-year grade level expectations for the previous grade. Students in this category should be ready to be taught the knowledge and skills inherent in the Common Core State Standards for their current grades.
   - Students in this category can be expected to expand their knowledge and skills mastery as the school year progresses. Table 54 includes three on-grade-level Scaled Score ranges for each grade that indicate progress that is approximately proportional to the time of year.
   - Students may require instructional intervention when they start the year with STAR Math scores that classify them "on grade level," but later fall behind the pace indicated for the fall, winter, or spring trimesters. Ideally, such

intervention should be preceded by a deeper assessment of strengths and weaknesses, and followed by ongoing progress monitoring.

Example: A student whose STAR Math FGL is “on grade level” appears to have mastered the knowledge and skills taught at lower grades, and should be expected to be ready for instruction at his/her current grade level.

2. **Below grade level.** Students in this category have Scaled Scores that indicate they have not attained the CCSS end-of-year grade level expectations for the previous grade. They may lack knowledge or skills that comprise one or more of the previous grade’s CCSS standards. A deeper assessment of knowledge and skills, followed by appropriate intervention, may be needed to bring these students up to grade level and prepare them to master the CCSS standards of their current grade.

   Example: A fifth-grade student with a STAR Math FGL classification of “below grade level” may have some gaps in fourth- or even lower-grade knowledge and skills. These gaps may need to be closed if the student is to fully benefit from instruction in fifth-grade topics.

3. **Above grade level.** Students in this category have Scaled Scores that indicate they have exceeded the minimum CCSS end-of-year grade level expectations for their current grade. A deeper assessment of knowledge and skills may be called for here, too, to verify their attainments of specific standards. Instruction of these students should probably aim at confirming or reinforcing their mastery of the current grade’s CCSS standards, expanding that mastery beyond the minimum end-of-year goals, and perhaps providing advanced-level instruction on skills one or more grade levels above the current grade.

   Example: A student whose STAR Math FGL is “above grade level” appears to have mastered the current grade’s knowledge and skills at the minimum 70% level, and perhaps more. This student may be ready to benefit from instruction in Math skills at a higher grade level. Teachers or administrators should verify this, based on additional assessment and their own knowledge of the student.

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. 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 Scaled Score of 868 correspond to a PR of 80, and, using the same grade placement, a STAR Math Scaled Score of 911 corresponds to a PR of 90. Thus, a difference of 43 Scaled Score points represents a 10-point difference in PR. However, for the same grade placement of 7.7, a STAR Math Scaled Score of 788 corresponds to a PR of 50, and a STAR Math Scaled Score of 812 corresponds to a PR of 60. While there is now only a 24-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 in "Normal Curve Equivalent (NCE)" on page 108, are much more appropriate for these types of calculations.

Table 58, "Scaled Score to Percentile Ranks Conversion by Grade (at Month 10 [June] in the School Year)," on page 129 contains an abridged version of the Scaled Score to Percentile Rank conversion table that is used for STAR Math. The actual table includes data for all of the monthly grade placement values from 1.0–12.9. Because the norming of STAR Math occurred in the seventh month of the 2002 school year, the seventh-month values for each grade are empirically based; these are the values in Table 58. The remaining monthly values were estimated by interpolating between the empirical points. The table also includes a column representing students who are just about to graduate from high school.

Table 58 can be used to estimate PR values for tests that were taken when the grade placement value of a student was incorrect (see "Grade Placement" on page 109 for more information). One always has the option of correcting the grade placement for the student if the error is caught right away, and then having the student retest. However, the correction technique using this table (illustrated in the example below) is intended to provide an alternate correction procedure that does not require retesting.

To illustrate, if a grade placement error occurred because a third-grade student who tested in April was accidentally entered as a fourth-grader, his or her Percentile Rank and NCE scores will be in considerable error. In order to obtain better estimates of this student's norm-referenced scores, look in the grade 3 column in Table 58 and locate the student's Scaled Score or the next-higher value in the table. Next, find the PR value associated with this particular Scaled Score for a student in month 7 of third grade. Then, follow the same procedure using the grade 4 column to obtain a PR corresponding to the same Scaled Score, had the student been in month 7 of fourth grade.

Teachers can use a similar interpolation procedure to obtain PR values that correspond to scores that would have been obtained at other times throughout the school year.

This procedure, however, is only an approximation technique designed to compensate for grossly incorrect scores that result from a student testing while his or her grade placement was incorrectly specified. A slightly better technique involves finding the PR values in Table 58 on page 129, converting them to NCE values using Table 59 on page 133, interpolating between the NCE values, and then converting the interpolated NCE value back to a PR value using Table 60 on page 134.

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 to 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 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 59 on page 133 provides the NCEs corresponding to integer PR values and facilitates the conversion of PRs to NCEs. Table 60 on page 134 provides the conversions from NCE to PR. The NCE values are given as a range of scores that convert to the corresponding PR value.

Student Growth Percentile (SGP)

Student Growth Percentiles (SGPs) are a norm-referenced quantification of individual student growth derived using quantile regression techniques. An SGP compares a student's growth to that of his or her academic peers nationwide. SGPs provide a measure of how a student changed from one STAR testing window[5] to the next relative to other students with similar starting STAR Math scores. SGPs range from 1–99 and interpretation is similar to that of Percentile Rank scores; lower numbers indicate lower relative growth and higher numbers show higher relative growth. For example, an SGP of 70 means that the student's growth from one test to another exceeds the growth of 70% of students nationwide in the same grade with a similar beginning (pretest) STAR Math score. All students, no matter their starting STAR score, have an equal chance to demonstrate growth at any of the 99 percentiles.

SGPs are often used to indicate whether a student's growth is more or less than can be expected. For example, without an SGP, a teacher would not know if a Scaled Score increase of 100 represents good, not-so-good, or average growth. This is because students of differing achievement levels in different grades grow at different rates relative to the STAR Math scale. For example, a high-achieving second-grader grows at a different rate than a low-achieving second-grader. Similarly, a high-achieving second-grader grows at a different rate than a high-achieving eighth-grader.

SGPs can be aggregated to describe typical growth for groups of students—for example, a class, grade, or school as a whole—by calculating the group's median, or middle, growth percentile. No matter how SGPs are aggregated, whether at the class, grade, or school level, the statistic and its interpretation remain the same. For example, if the students in one class have a median SGP of 62, that particular group of students, on average, achieved higher growth than their academic peers.

Grade Placement

It is very important that the STAR Math software uses students' correct grade placement values when determining norm-referenced scores. The values of PR (Percentile Rank) and NCE (Normal Curve Equivalent) are based not only on what Scaled Score the student achieved, but also on the grade placement of the student at

5. We collect data for our growth norms during three different time periods: fall, winter, and spring. More information about these time periods is provided on page 119.

the time of the test. For example, a second-grader in the seventh month with a Scaled Score of 534 would have a PR of 65, while a third-grader in the seventh month with the same Scaled Score would have a PR of 24.

Thus, it is crucial that the STAR Math software contains the proper grade placement, and that any testing in July or August reflects the proper understanding of how STAR Math deals with these months in determining grade placement, described below.

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 or Math Instructional Level (MIL) using whole numbers, the STAR Math software automatically adds fractional increments to that grade based on the month of the test. To determine the appropriate increment, STAR Math considers the standard school year to run from September–June and assigns increment values of 0.0–0.9 to these months. 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 Place RT 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 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 50, "Incremental Grade Placement Values per Month," on page 101 summarizes the increment values assigned to each month.

If your school follows the standard school calendar used in STAR Math 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 in the *Renaissance Place Real Time Software Manual*.

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. The STAR Math 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. If you discover that a student has tested with an incorrect grade placement assignment (use the Growth, Screening, Summary, or Test Record reports to find out the grade placement), the procedures outlined in the last paragraph under "Percentile Rank (PR)" on page 107 (in the discussion about Table 58) can be used to arrive at corrected estimates for the student's Percentile Rank and Normal Curve Equivalent scores.

STAR Math in the Classroom

There are numerous ways that STAR Math can be used in the classroom, as well as at the school and district level. At the classroom, grade, school, or district level, it can be a useful tool for instructional planning, growth measurement, and program evaluation. At the individual level, it can be used for a variety of purposes, including screening, formative assessment, progress monitoring, and outcomes assessment. This section provides examples of how to use STAR Math for many of these purposes.

Goal Setting and Instructional Planning

Goal setting is an almost ubiquitous practice in education. Teachers continually set goals for their students, and administrators set goals for their schools. By setting clear and achievable goals people are able to comport their behavior in an appropriate manner towards achieving those goals. This is true of school-wide or classroom-specific goals. However, not all goals are set equally. Some goals may be set ambiguously or lack a clear and measurable frame of reference. Good goal setting includes setting realistic and measurable goals that are achievable within the time frame identified.

Goals can provide a clear set of expectations of what must be accomplished and in what amount of time. It is also possible to break down long-term goals into a series of intermediate objectives or short-term goals. This can help to focus time and energy on the important aspects of meeting the long-term goal at shorter and more manageable increments. It also provides a standard for which a person may strive. Goals can also be motivating in that the realization of them provides a sense of accomplishment and achievement.

There are a few crucial aspects to goal setting in general. One of the essential aspects of goal setting is to set a measurable goal objective for some point in the future. This goal must be measurable so as to establish a criterion that represents accomplishment. It is also useful to set a series of intermediate, measurable steps to accomplish that goal. This provides a method of incremental evaluation of the progress being made towards the long-term goal. The power of this method is that it can provide early warning signals with respect to potential problems meeting the goal or recognition that one is on-track to meeting the stated objective in the future. These types of signals are important for an objective evaluation of progress. This is one of the main reasons educators need reliable and valid measurements.

If we are to measure progress and goal attainment, we need to be sure that the measuring device actually measures what we think it measures, and that it does so consistently. As an extreme example, if our long-term goal was to have our students master fractions and we used a reading test to measure progress, we should not be surprised when the signals we receive from a reading test provide no relevant information on development towards rational numbers.

It is also important that the assessment measure we use provides consistent scores, because we would like to be confident that the score a student received actually tells us with a high level of precision what the student's actual ability level is.

STAR Math provides a reliable and valid method for measuring progress towards achievable goals in mathematics. By using STAR Math on a regular basis, such as quarterly or monthly, teachers can monitor students' progress and make appropriate adjustments to instructional practices. Progress monitoring is an approach that has strong research support and has proven successful in a variety of educational settings.

STAR Math also provides practical advantages over other methods of gathering multiple pieces of data over time needed for monitoring achievement towards a set goal. It takes only about 10–15 minutes to administer (20 minutes for STAR Math Enterprise); its brief administration time helps maximize the amount of in-class time available for instruction. Results are also provided immediately to the teacher so the teacher will be able to review the student's progress more quickly than with most assessments.

STAR Math can also be administered at different times and with different frequencies for different students. This allows the teacher to specify and make professional decisions based on intermediate assessments, on a student-by-student basis. It also allows the teacher to measure a student's specific response to any type of intervention being provided. This helps to strengthen the teacher's ability to make real-time, professional decisions about instructional approaches for each student.

STAR Math can also be administered quite frequently. This allows the results of the assessment to be graphed in order to show growth. Charting progress in this way can be used both at the individual and classroom level as an evaluative check to monitor effectiveness. Periodic charting of progress can also be motivating, as students visualize their progress and recognize their achievement. This type of ongoing information gathering can be used for a variety of different functions within a school; examples include parent-teacher meetings and child-study team meetings where groups of teachers discuss ways to intervene with struggling students.

The STAR Math assessment also has been shown to be highly related to state assessments and widely used standardized tests. This can facilitate critical benchmarking of student achievement across the grades. STAR Math does not specifically measure states' instructional standards, but scores on STAR Math assessments are statistically related to those proficiency standards. Therefore, scores on STAR Math can be used to predict later outcomes. This type of information is useful in forecasting educational achievement and making decisions about utilizing resources with respect to a student's instruction. It is also possible to employ more complex school- or district-wide implementations of the assessment to gauge student progress towards the all-important end-of-year goals consistent with a state's educational standards.

To interpret screening results, schools often use benchmarks and cut scores. These scores help educators identify which students require some form of intervention to accelerate growth and move toward proficiency. A goal-setting wizard is used in the program to set and track goals; the Screening Report and the Student Progress Monitoring Report are used to track students' progress towards goals and growth. (See the *STAR Math Software Manual* for more information.)

Formative Assessment

The purpose of formative assessment is to improve student learning by providing the teacher with relevant information for instruction. STAR Math accomplishes this purpose by providing the teacher with valid and reliable information regarding the current achievement of students in mathematics. STAR Math is sensitive to small changes in math skills, and it has a high upper range so there is no ceiling effect for most grades.

Measuring Growth

When evaluating or assessing the academic and educational achievement of students, it is important to estimate the amount of growth students obtain within a school year and also across multiple school years. There are many problems inherent in measuring growth from conventional paper-and-pencil tests within a grade and even more problems associated with measuring growth across multiple grades (see Kolen & Brennan, 2004 for more in-depth discussion). STAR Math addresses these problems by using a technique called vertical scaling, which allows all students' scores to be placed on the same developmental score scale. This provides score comparability within a school year and allows students or cohorts to be followed across multiple school years.

Absolute versus Relative Growth

It is important to distinguish between two types of academic growth (or gains) that may be evidenced in test results: absolute growth and relative growth. Absolute growth reflects any and all growth that has occurred. For example, as a child matures, we can see absolute growth in his height, as measured in feet and inches or meters. Relative growth reflects only growth that is above and beyond "normal" growth (i.e., beyond typical growth in a reference or norming group). This measure of growth identifies a student's growth or gains relative to a reference group of students over the same or similar period of time.

As an example, imagine a group of students whose test results place them at the 40th percentile, with an average Scaled Score of 686, in the spring of grade 5. In the spring of grade 6, the same group still scores at the 40th percentile with an average Scaled Score of 737. This group of students has experienced 51 Scaled Score points of absolute growth, but there has been no relative growth (since the group scored at the 40th percentile in both grade distributions). In other words, relative growth will only be positive when growth has exceeded "normal" growth as defined by the norming or reference sample. In general, norm-referenced scores such as percentiles and NCE scores only indicate relative growth, whereas Scaled Scores (and Grade Equivalent scores) reflect absolute growth. The STAR Math Growth Report provides you with information about both aspects of growth. In general, most educational program evaluation designs attempt to determine whether relative growth has occurred. That is, they are attempting to measure the impact of the intervention, or program, above and beyond normal growth.

Methods of Measuring Growth

New interventions are continually being proposed for educational settings, most with the aim of improving educational outcomes. Such interventions may be extensive, such as a new teaching method or new curriculum, or they may be smaller in scope, such as a new textbook. The introduction of a Tier 1 progress monitoring system, such as Accelerated Math, into a school or classroom is a good example of such an intervention. Whatever the proposed intervention, however, it is first necessary to establish its effectiveness in terms of the educational benefit for students. Examination of the effectiveness of new teaching methods, a new curriculum, and other such interventions is extremely important if we are to accurately determine whether these programs and/or methods are working. This is important for appropriate direction of limited resources and for ensuring that those programs, which will have the most educational impact on children are clearly identified.

Along with identifying whether or not an intervention is effective by use of a final summative evaluation, ongoing formative evaluations are also important. The evaluation of student progress is an ongoing procedure as the students learn and apply principles and facts learned in the classroom to solve everyday problems. Therefore, the measurement of growth can be seen as a descriptive method for understanding the developmental path of students as they acquire certain skills and enhance other abilities. With the use of on going monitoring of progress, teachers may be able to intervene more quickly to alter the course of instruction for a group or even more specifically to an intervention targeted at one or a few students who may be struggling. However, the monitoring of progress on an individual or small-group basis is not limited to only students with high needs, but can also be used to monitor the progress of high-achieving students who may be provided more free time to explore individual interests.

The measurement of growth is a long-established tradition in social sciences in general and education specifically. While this is a large and important area of exploration, the depth of methodological and statistical analysis available at present cannot be fully described in a technical manual. The intention of the following sections is to provide a general overview of possible methods of evaluating growth using STAR Math. We will also provide a list of reference materials at the close of the manual for interested readers to pursue a more thorough investigation of current methods of analysis and design.

Pretest/Posttest Designs

One of the simplest methods for evaluating the effect of an intervention is the pretest-posttest paradigm, in which students are assessed twice—once prior to intervention, and once again at its completion. This method was born out of the experimental methodologies of science in an effort to quantify changes in an outcome variable by isolating the independent variables in a given system. For instance, if one would like to know whether a specific intervention increases multiplication skills or phonemic segmentation, one would isolate a sample of students, randomly assign half of the students to a no-intervention group and the other half to intervention, and assess all of them before and after the intervention. Then one would look for differences in outcomes between the two groups, assuming the intervention is the only systematic difference between the groups, and make a claim about whether or not the students in the intervention group did better when compared to the students who did not receive the intervention (the no-intervention model).

An experiment with a pretest/posttest design can utilize a control group of students, who like the above example, do not receive the intervention. This provides a comparison group against which to gauge the practical effects of the intervention applied to the intervention or treatment group and make inferences about intervention effectiveness over and above those without the intervention.

However, sometimes the use of a control group is not feasible. Under these circumstances, educators may opt to utilize norm-referenced scores, such as percentile ranks or normal curve equivalents (NCE) scores. For example, a school may introduce a new curriculum to a whole grade level and thus would not have a readily available control group. The school may decide to use a "proxy" comparison group by utilizing norm-referenced scores. In effect, the test developer's norming group is being used as a proxy for a control group who are not provided the intervention. This allows changes in relative growth to be evaluated against the norming group.

In such a design, each student is administered a test prior to the beginning of the intervention to establish a baseline measure. Then, each student is measured again at a later point in time (usually with a different, but equated, "form" of the same test) to see whether the intervention is providing the desired outcome. The follow-up measurement may be at the end of the intervention, or may be done periodically throughout the course of the new program. Certainly, all of the issues relating to the technical adequacy of the test itself (e.g., reliability and validity) are applicable in order for this type of research to work properly. One key factor in conducting pretest/posttest designs is that if the same test form is used both times, then the results may be compromised due to students having previously been exposed to the test items. In an ideal situation, equivalent tests with no items in common should be administered; STAR Math is ideal for this, because tests administered to a student within 75 days of one another will have no items in common.

When the test scores used in the evaluation are norm-referenced (such as Percentile Ranks), a control group is not necessarily required since the scores themselves allow you to compare growth to that of the peer (norming) group. It should be noted that when a test is normed, the percentile information is derived based on the specific point during the academic year when the test was administered. For example, suppose a test was normed in the spring (seven months into the school year) but a teacher wants to make an assessment at the beginning of the school year.

In order to provide normative information for each month of the academic year, STAR Math software examines the difference between adjacent grade levels and, presuming even growth, interpolates between the empirical (observed) norms. Caution should be exercised when looking at growth that is based on these interpolated percentiles. This is because the assumption that growth occurs evenly over the time period (i.e., between the adjacent empirical percentiles) may be unrealistic.

The goal of this type of study is to determine if a program intervention has resulted in improvement beyond what is expected based on the norming population (i.e., to see if the posttest results place the students above where they would be if there had not been any intervention). For example, if a group of 4th-grade students' pretest scores indicate that their group percentile (corresponding to the average NCE) is 25, then we want to see if their 5th-grade posttest scores will result in a group percentile that is greater than 25. (Caution must be exercised in cases where average pretest scores are substantially above or below the norm, however. Due to the phenomenon known as “regression to the mean,” posttest scores will tend to move towards the norms group mean even if no real change has occurred. Consequently, corrections for regression to the mean may need to be applied before the results of an experimental intervention are interpreted.)

When comparing the students’ growth to growth based on norms, only one group is required, but in this case, the time period between pretest and posttest should be at least one year; otherwise the growth would be referenced against interpolated data. This corresponds with US Department of Education recommendations for Chapter I (Title I) program impact studies, which state that:

> The general rule of thumb for norm-referenced evaluations is that testing should be done within two weeks of the midpoint of the empirical norming period *(U.S.D.E. Evaluator’s References for Title I Evaluation and Reporting System, Volume 2)*.

For the STAR Math test, the empirical norming period was in the month of April 2002. The US Department of Education further recommends that interpolated norms that vary by more than six weeks from the empirical data points should not be used for norm-referenced evaluations. In general, a good rule of thumb regarding sample size requirements for any growth study is “more is better!” As the size of the group increases, you can be more confident that the obtained results are genuine.

The construction of STAR Math ensures that students get psychometrically parallel versions of the test at both pretest and posttest administrations. Thus student growth can be directly measured without any confounding problems related to having seen items at the previous time of measurement. It is important to note that growth is best measured at a group level, such as a classroom or grade level. This is because at the individual student level, there are technical issues of unreliability associated with growth (gain) scores, and measurement error causes fluctuations of individual students’ scale scores that could mask the true amount of growth.

Longitudinal Designs

Longitudinal designs are different from pretest/posttest designs in that data is gathered on the same students multiple times over an extended time period. Some people argue that the evaluation of only two time points like the pretest/posttest design does not successfully identify a longitudinal design.

A longitudinal design has a least three time points of measurement. An example of this approach can be seen in the assessment of students in the fall, winter, and spring quarters of the school year.

The basis for the longitudinal design is to gather ongoing information on student development. This allows for an identification of trends in student achievement along with normal developmental trends with which to compare student growth. Usually, one is interested in how students change over a period of time and finds this change as an indication of instructional and/or intervention efficacy.

Longitudinal designs are very useful as formative evaluations but can also be used in conjunction with summative evaluations. For example, a goal level may be specified for an end-of-the-year evaluation. This would be the summative feature that endeavors to evaluate whether or not the goal was obtained in the time period designated. However, one can incorporate a longitudinal design by more frequently measuring student progress, e.g., at quarterly or monthly intervals. This would allow a teacher to track progress on a monthly basis as the classroom moves towards the stated end-of-year goal. This is also very informative as it provides a signaling system for the teacher if the students begin to fall behind or are not progressing at an expected pace.

Student Growth Percentile (SGP)

Because STAR Math is so widely used, Renaissance Learning has data for millions of testing events. With these scores, we are able to calculate growth norms. In other words, we can approximate how much growth is typical for students of different achievement levels in different grades from one time period to another. Renaissance Learning first incorporated growth modeling into STAR Math reporting in 2008 via decile-based growth norms. Student Growth Percentiles (SGPs) represent the latest advancement in helping educators understand student growth. SGPs are available in STAR Math for grades 1–12.

SGPs are a normative quantification of individual student growth derived using quantile regression techniques. An SGP compares a student's growth to that of his or her academic peers nationwide. SGPs provide a measure of how a student changed from one STAR testing window[6] to the next relative to other students with similar starting STAR Math scores. SGPs range from 1–99 and interpretation is similar to that of Percentile Rank scores; lower numbers indicate lower relative growth and higher numbers show higher relative growth. For example, an SGP of 70 means that the student's growth from one test to another exceeds the growth of 70% of students in the same grade with a similar beginning (pretest) STAR Math score.

SGP was initially developed for use in state summative assessments (Betebenner, 2010). The STAR assessments are believed to be the first non-state assessment and the first interim test to incorporate SGP. SGP has currently been adopted by 11 states for use with their summative tests. Those states are Colorado, Idaho, Indiana, Massachusetts, Nevada, New Hampshire, New Mexico, Rhode Island, Virginia, West Virginia, and New Jersey. In applying the SGP approach to STAR data, Renaissance Learning has worked closely with the lead developer of SGP, Dr. Damian Betebenner, of the Center for Assessment, as well as technical advisor Dr. Daniel Bolt, an expert in quantitative methods and educational measurement from the University of Wisconsin–Madison.

Applying the SGP approach to interim assessment data involved a number of technical challenges, primarily the differences regarding how STAR Math and state tests are administered. State summative tests are typically administered once a year, at approximately the same time, to all students. On the other hand, STAR Math is much more flexible, and may be administered to students as often as weekly. Decisions on when to administer and which students will participate are left to local educators. Most commonly, schools use STAR Math as a screening and benchmarking test for all or nearly all students 2–4 times per year. Students requiring more frequent progress

6. We collect data for our growth norms during three different time periods: fall, winter, and spring. More information about these time periods is provided later in this section.

monitoring may take STAR Math on a more frequent basis to inform instructional decisions, such as whether the student is responding adequately to an intervention. Because of this flexibility, not all students necessarily take STAR Math at the same time; the number and dates of administration vary from one student to the next. However, the majority of students test within at least two of the following time periods: fall (August 1–September 30), winter (December 1–January 31), and/or spring (April 1–May 31). We chose these date ranges when defining the data sets that would be used to determine Student Growth Percentiles. Therefore, we can provide Student Growth Percentiles for achievement that takes place between fall and winter STAR testing, winter and spring STAR testing, and/or fall and spring STAR testing, as defined above.

To calculate Student Growth Percentiles, Renaissance Learning collected hosted student data from the five most recent school years (2006–07, 2007–08, 2008–09, 2009–10, and 2010–11). Sample sizes were approximately 2 million students. Table 55 has details on demographics of these students. Quantile regression was applied to characterize the bivariate distribution of students' initial scores and ending scores. Students were grouped by grade and subject, and then quantile regression was used to associate with every possible initial score and ending score combination a percentile corresponding to the conditional distribution of end score given the initial score. The result of these analyses was the creation of a look-up table in which beginning and ending student STAR scores are used as input to define a student growth percentile for each grade, subject, and time period (e.g., fall to winter, winter to spring, fall to spring).[7] The use of quantile regression techniques makes construction of such tables possible even though not all possible initial and ending score combinations were actually observed in the student data. Loosely speaking, the quantile regression approach can be viewed as a type of smoothing in which information from neighboring score values (initial and ending) can be used to inform percentiles for hypothetical score combinations not yet observed. As such, application of the methodology allows us to look up any score combination to obtain the percentile cutpoints for ending score conditional achievement distribution associated with the given initial score. These cutpoints are the percentiles of the conditional distribution associated with the student's prior achievement. Specifically, using the quantile regression results of grade 6 STAR Math spring scores on fall scores, estimation of the 1st, 2nd, 3rd,…99th percentile growth from fall to spring can be calculated. Using each of these cutpoints, we are able to calculate a Student Growth Percentile for every subject, grade, and score combination.[8] (Betebenner, 2009, Betebenner, 2011a & Betebenner, 2011b).

7. Because we use a national baseline for calculating SGPs, students are compared to their peers nationwide based on prior years' data, rather than their peers within the same school, district or state. This, combined with our use of a static lookup table, ensures that every student with the same pre- and post-test STAR Math score combination will have the same SGP no matter how many other students have that same combination. There is no cap on the number of students receiving a particular SGP.
8. Expert recommendation was that we should not report Student Growth Percentiles for extremely unusual pretest scores. Therefore, we do not report Student Growth Percentiles for those students with an extremely low or high pretest score.

Table 55: Sample Characteristics, STAR Math SGP Study

		Sample %		
		Fall to Spring (n = 2,316,561)	Fall to Winter (n = 1,680,004)	Winter to Spring (n = 1,679,303)
Geographic Region	Midwest	24.2%	25.7%	27.4%
	Northeast	4.0%	4.0%	4.9%
	South	51.3%	52.5%	49.7%
	West	20.6%	17.7%	18.0%
	Response Rate	95.7%	95.2%	94.9%
School Type	Public	95.4%	95.6%	95.6%
	Private, Catholic	1.7%	2.8%	2.8%
	Private, Other	2.9%	1.6%	1.6%
	Response Rate	90.8%	89.9%	89.0%
School Enrollment	< 200	5.4%	5.4%	5.5%
	200–499	41.5%	42.5%	42.8%
	500–2,499	52.6%	51.8%	51.4%
	2,500+	0.4%	0.4%	0.3%
	Response Rate	93.0%	92.2%	91.4%
School Location	Urban	20.4%	19.3%	21.7%
	Suburban	26.2%	23.7%	25.8%
	Rural	53.3%	48.3%	52.4%
	Response Rate	92.1%	91.4%	90.6%
Ethnic Group	Asian	3.5%	3.8%	3.7%
	Black	20.6%	21.5%	22.6%
	Hispanic	18.7%	17.4%	17.8%
	Native American	4.5%	4.8%	4.7%
	White	52.7%	52.5%	51.2%
	Response Rate	36.5%	35.4%	35.2%
Gender	Female	49.2%	49.0%	49.1%
	Male	50.8%	51.0%	50.9%
	Response Rate	70.5%	69.5%	69.8%

Periodic Improvement

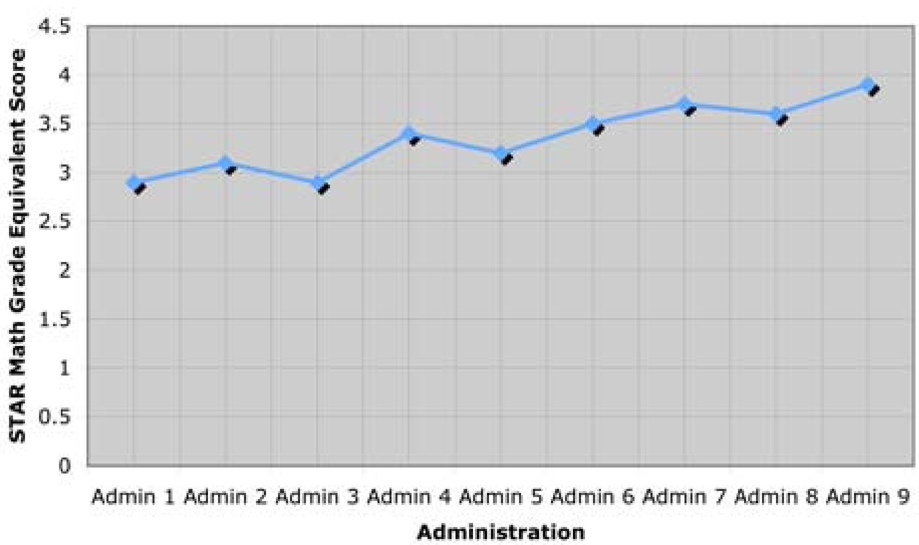
The Grade Equivalent Score can be used for measuring periodic improvement because it is reported in tenths of a grade. The correspondence between decimal value and month is shown in Table 56.

Table 56: Correspondence between Month and Decimal Value

Month	Decimal Equivalent	Month	Decimal Equivalent
September	0.0	February	0.5
October	0.1	March	0.6
November	0.2	April	0.7
December	0.3	May	0.8
January	0.4	June	0.9

The Grade Equivalent score generated by STAR Math makes it possible to track the progress students should make on a monthly and annual basis. It is important to keep in mind, however, that the month-to-month Grade Equivalent Scores for a student are unlikely to move upward consistently. Students making appropriate progress may nonetheless show an erratic growth trajectory. Figure 5 on page 121 shows the score trajectory for a typical third-grade student for nine monthly administrations of STAR Math.

Figure 5: Monthly Progress of a Third Grader



The student started the year a little below the 3.0 GE at approximately a GE of 2.9 and is showing approximately a year's growth from initial to final assessments, but the trajectory of growth was erratic. This growth pattern is to be expected and reflects the measurement error in tests and the fluctuation in students' test performance from one occasion to another.

A decline in Grade Equivalent Score from one test to the next is not a matter of concern unless it persists for two or more assessments. Intermittent score declines and erratic trajectories are not unique to STAR Math. They happen with all other tests that are administered at frequent intervals. A good example of this is the progress graph reported in “Developments in Curriculum-Based Measurement” (Deno, 2003).

In conclusion, STAR Math provides an efficient and useful measure of growth for both formative and summative evaluations using both pretest/posttest and longitudinal designs. STAR Math addresses many of the problems normally associated with measuring growth over time. One of those is the time involved in assessing multiple students many times throughout the year. With STAR Math, each student can take the assessment in about 10–15 minutes (20 minutes for STAR Math Enterprise) and at any time during the monthly period. Therefore, using STAR Math, the teacher can maximize instructional time for the class as a whole and minimize the assessment time for each student. Also, since the scoring is done automatically, the teacher is able to receive rapid feedback without the time associated with scoring each student’s assessment protocol.

In the context of progress monitoring, RTI and periodic improvement methods, STAR Math provides a reliable and valid, norm-referenced measure of a student’s math achievement. This can be used to establish a baseline measure of student ability and to evaluate student growth over time. This type of information is vital since many times in the educational setting one is unable to define a control or reference group to which one will make later comparisons.

Growth Estimates

One important aspect of measuring growth is to have a standard by which to evaluate it. For instance, if someone told you a student gained 25 scale score points in a year, how would you be able to evaluate it and make a judgment about how well the student is developing? It would be almost impossible without a frame of reference to evaluate the extent to which the student profited from instruction. Therefore, it is important to have some way of interpreting the test score growth a student exhibits. One useful method of doing this would be to relate a student’s growth to an estimate of what would be normal growth for a similar student.

With an estimate of expected growth for a student based on growth estimates of similar students, one would then be able to make statements as to whether or not a student made the growth expected within the specific time frame. For instance, many schools and districts use STAR Math to measure students at the beginning, middle, and the end of the school year to evaluate how much the school has contributed to the students’ learning. Other schools and districts use STAR Math as a summative assessment towards the end of the school year and then use that to gauge growth by the next school year at the same time. Also, now that schools are subject to state accountability regulations in compliance with the No Child Left Behind Act of 2001 (NCLB), some schools administer a screening assessment at the beginning of the school year to identify students believed to be at risk of failing to meet the later math standards, and then administer follow-up tests to monitor the progress of these students throughout the school year. STAR Math is highly useful for these screening and progress-monitoring functions, given its efficiency, ease of use, and excellent technical qualities.

STAR Math's vertically scaled test scores (Scaled Scores) allow student scores to be compared across grades as well as within grades. When comparing the growth of students, it is important to have some idea of how much they should be growing normally to evaluate whether or not a program actually increased the growth of a student. Without an expected growth estimate, teachers and administrators may make invalid inferences about the value of a program simply because of normal maturation over time.

In evaluating growth over time, it is important to take grade levels of students into consideration. Two students at different grade levels who attain the same Scaled Score on STAR Math may have dramatically different expected growth scores over the same period of time. For instance, suppose a first grader and a second grader both obtain scale scores of 425 on an assessment taken during April of the same school year. It would be wrong to assume that they both should grow the same amount.

In fact, a student scoring 425 at the end of first grade would be expected to obtain a scale score of about 534 by the end of the next school year, while the second grader would only be expected to score around a 492 the next school year.

Growth is different for different age groups and also different within an age group depending on where students fall in the distribution of abilities. For instance, take the first-grade student who scored 425 at the end of the year. This student was expected to score about 109 scale score units higher by the same time in the following school year. However, a similar aged student in the first grade who scored 269 at the same time would be expected to have a score around 400 by the same time during the next school year. This student is expected to grow by 131 scale score units. Therefore, a single estimate of growth even within a grade can be highly misleading.

To estimate the normal amount of growth from year to year, one must take into account both the grade level of the student at the time of the initial evaluation and also the performance level of the student. To facilitate the use of STAR Math scores for estimating growth for students, one can use the normative data, or one can use empirical data derived from one's own district or school. The use of empirical support for making estimates about growth will be developed in the following section with examples. For the rest of this section we will outline a basic method using the normative data.

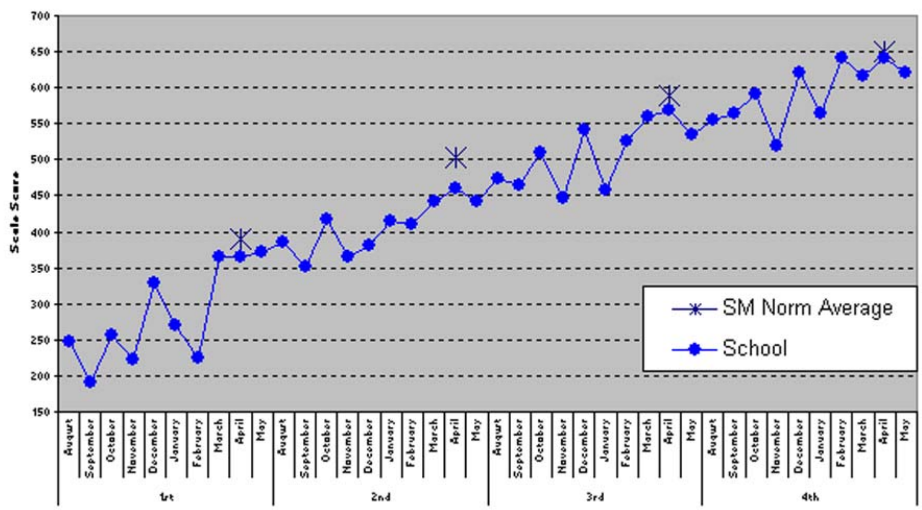
Growth Measurement Examples

To provide a basic example of how a school or district can use STAR Math as a growth measure, we will use some actual data from a school that is serving students from kindergarten to fourth grade. The school tests all students on a monthly basis using STAR Math. We will elaborate on two important uses of growth measurements. The first has to do with a group of students aggregated together, such as a school, a grade, a grade within a school, or even a classroom. The other example is relevant to progress monitoring for an individual student.

Growth Measurement at the Group Level

The school can graph the monthly averages at each grade in a manner similar to Figure 6. Here the school is computing the average scale score for students at each month and plotting them on a graph. They are also plotting the STAR Math norm data provided in the technical manual for each April of the school year. This provides a reference point for the teachers and administrators to gauge how well their students are progressing relative to STAR Math national norms.

Figure 6: Average STAR Math Scaled Scores for the School and the Norms Group



There are many ways to extend the use of the data to help in making informed decisions. For instance, the school can break down the averages for each month by classroom for comparative purposes.

Another way is to use data from the school to predict later outcomes, using statistical prediction models such as linear regression. A driving reason behind developing predictive models is the present educational accountability standards in each state. It would be very useful to be able to predict a student's likely performance on state tests and then provide some type of intervention early in the year if the student appears to be at-risk of not attaining the proficiency standard.

For this example, we will predict end-of-year STAR Math scores from the beginning of the year scores. This will illustrate a basic methodological approach that can be extended to predicting student scores on a state-mandated accountability test quite easily. For purposes of illustration, the example is limited to fourth-grade students. We choose to focus on a single grade because this would probably be the preferred method when trying to predict a student's outcome on a state test of proficiency that is aligned to grade-level standards.

The easiest and most basic method to predict later outcomes is with the use of a simple regression model (Neter, Kutner, Nachtsheim, & Wasserman, 1996). In this situation the outcome variable would be the end of year STAR Math scale score, and the predictor variable would be the beginning of the year STAR Math scale score. For this example, we will use the average of the August and September scale scores as the predictor variable and will use the average of the April and May scale scores for the outcome. (Averaging scores in this way increases the reliability of the variables and hence the accuracy of the predictive models. Single data points may be used if a school or district only used STAR Math to assess math achievement at three times during the year like the fall, winter, and spring sessions.)

Analysis of STAR Math data from an empirical database indicates that the best fitting prediction line is

End-of-Year Scale Score = 66 + 1.03 (Beginning-of-Year Scale Score).

This model then allows for the prediction of student end-of-year scores from their beginning-of-year scores. For instance, if a student scored a 550 at the beginning of the year, we'd expect that the student would score about 633 at the end of the year. This not only gives us an idea of how much a student should gain over a year in a particular grade, it can also directly inform progress monitoring applications (see "Progress Monitoring" on page 125) by providing a standard about which to judge student progress.

This same approach can be applied to a state's end-of-year test of accountability in either of two ways. First of all, a school or district can give STAR Math assessments around the same time as the state test at the end of the year. Then the student scores on STAR Math and the state math test can be linked (Kolen & Brennan, 2004). From this linking, one can estimate the STAR Math scale score that is approximately equivalent to the threshold score for proficiency. Using this score, teachers and administrators can then set a STAR Math goal for each grade. Teachers can then also use individual students' beginning of the year scores to predict their end-of-year scores and then identify students who are not likely to make the proficiency benchmark score without intervention. This provides a method for intervening early in the year on behalf of students who appear to require special instruction in necessary skills to reach end-of-year proficiency.

A second method would be to use the state test of accountability score as the outcome variable and the beginning of the year STAR Math scale score as the predictor. This would allow administrators and teachers to predict end-of-year state test scores based on their beginning of the year STAR Math scores.

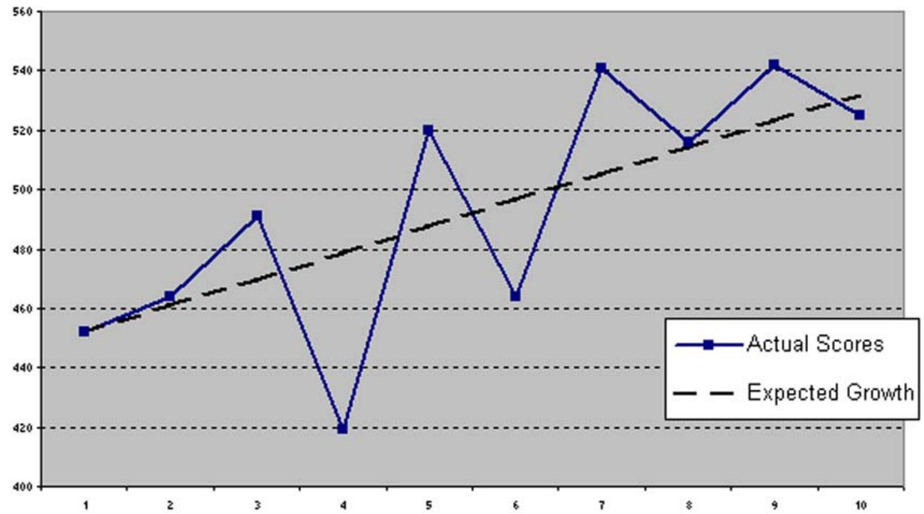
Progress Monitoring

Beginning in March 2008, Renaissance Place RT editions include Annual Progress reports. These reports contain graphical displays of individual and class scores that include STAR Math scores from all tests administered within the current school year. Using these reports, teachers can compare students' progress with that of a national norms group of students in the same grade.

Because the report is not available in prior versions of STAR Math, an example of plotting an individual's progress manually is described in the following text. Either the report or the plot can be used for the intensive progress monitoring of a student who may be perceived to be at risk or simply falling behind in the subject matter.

To extend the analysis above, we use a fourth-grade student. Using the equation in the previous section, we can predict a student's end-of-year score from that student's beginning-of-year score. Our example student has a scale score of 452 at the beginning of the year. Using the above predictive equation, we would expect the student to get a scale score of 532 by the end of the year. In Figure 7, we see a graph of the student's scores at monthly intervals along with a straight line showing the expected growth of the student over the year. This expected growth line is computed simply by taking the beginning-of-year score and end-of-year expected score and connecting them with a straight line.

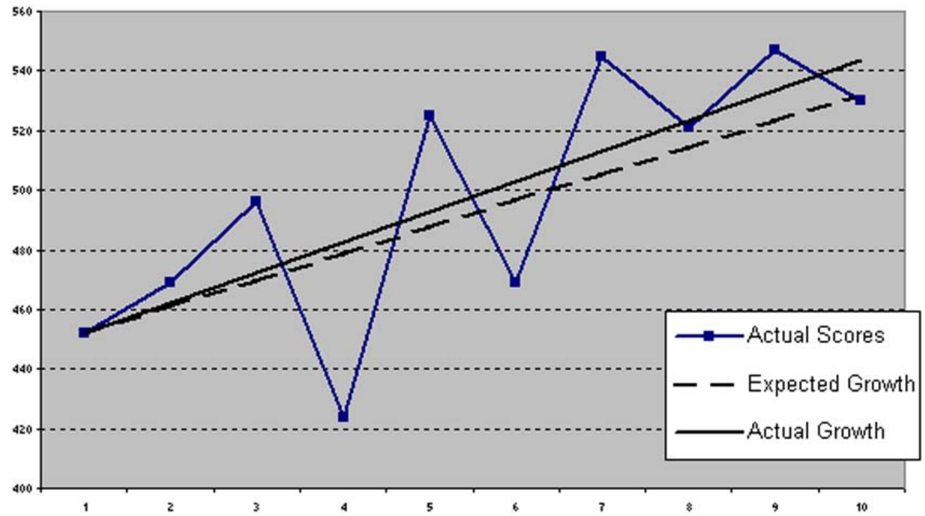
Figure 7: A Progress Plot of One Fourth-Grade Student's Actual STAR Math Scale Scores by Month, Compared to the Progress Predicted Using Linear Regression



The graph indicates that the student was on track across the year to meet the expected end-of-year goal. However, if the student was a lower-performing student, and interventions were provided throughout the year to help remediate the student's skills, then we would expect the student's actual growth to be different from the expected growth. For example, imagine the above student was provided an intensive intervention to improve math skills. Now suppose that the intervention increased the student's math ability by a modest five scale score points per month. In this case, the student will still have the same expected growth rate based on the beginning of the year score, but his/her actual trajectory should be higher than the "expected" one in response to the intervention.

Figure 8 shows this new situation along with an additional trend line for the student. This trend line is based on the student's monthly scale scores and provides an estimate of the student's actual growth over the year. We can see that the trend line for the student increases more rapidly than the "expected" growth trend. This provides a method to evaluate whether or not an intervention or series of interventions are having any measurable effect on the student's academic achievement. It is possible to do this same type of analysis on a more compact schedule rather than waiting until the end of the year. For instance, the goal might be the end of a semester and the intermediate measures done monthly.

Figure 8: New Situation with Additional Student Trend



STAR Math and No Child Left Behind

STAR Math may be useful for districts and schools as they conform to the 2001 No Child Left Behind legislation. For example, No Child Left Behind required states, starting in 2005, to annually measure the mathematics progress of students in grades 3–8. As noted throughout this manual, STAR Math is a reliable and valid measure of math achievement for students in grades 1–12. Furthermore, due to its computer-adaptive features, STAR Math requires less administration time and supervision than paper-and-pencil tests without compromising the psychometric quality of scores.

No Child Left Behind also requires that federal funding go only to those math programs that are backed by scientific evidence. As noted in the above section on growth measurement, teachers and administrators can use STAR Math to evaluate the effectiveness of math programs and interventions. Given the increased emphasis being placed on using only research-based teaching methods, more and more teachers will find STAR Math an invaluable tool in the process of demonstrating growth in mathematics achievement resulting from their math programs.

Conversion Tables

Table 57: Scaled Score to Grade Equivalent Conversions

Scaled Score	Grade Equivalent[a]	Scaled Score	Grade Equivalent	Scaled Score	Grade Equivalent	Scaled Score	Grade Equivalent	Scaled Score	Grade Equivalent[b]
0–163	0.0	482–491	2.7	691–695	5.4	794–795	8.1	825–826	10.8
164–176	0.1	492–501	2.8	696–701	5.5	796–797	8.2	827	10.9
177–189	0.2	502–510	2.9	702–706	5.6	798–799	8.3	828	11.0
190–202	0.3	511–520	3.0	707–711	5.7	800–801	8.4	829–830	11.1
203–215	0.4	521–529	3.1	712–716	5.8	802–803	8.5	831	11.2
216–228	0.5	530–538	3.2	717–721	5.9	804–805	8.6	832	11.3
229–241	0.6	539–547	3.3	722–726	6.0	806	8.7	833	11.4
242–254	0.7	548–556	3.4	727–730	6.1	807–808	8.8	834–835	11.5
255–267	0.8	557–564	3.5	731–735	6.2	809	8.9	836	11.6
268–280	0.9	565–573	3.6	736–739	6.3	810–811	9.0	837–838	11.7
281–294	1.0	574–581	3.7	740–743	6.4	812	9.1	839	11.8
295–307	1.1	582–589	3.8	744–747	6.5	813	9.2	840–841	11.9
308–320	1.2	590–597	3.9	748–751	6.6	814	9.3	842	12.0
321–333	1.3	598–604	4.0	752–755	6.7	815	9.4	843–844	12.1
334–345	1.4	605–612	4.1	756–758	6.8	816	9.5	845	12.2
346–358	1.5	613–619	4.2	759–762	6.9	817	9.6	846–847	12.3
359–370	1.6	620–627	4.3	763–765	7.0	818	9.7	848	12.4
371–382	1.7	628–634	4.4	766–768	7.1	819	9.8	849–850	12.5
383–394	1.8	635–641	4.5	769–772	7.2		9.9	851	12.6
395–405	1.9	642–647	4.6	773–775	7.3	820	10.0	852–853	12.7
406–417	2.0	648–654	4.7	776–778	7.4	821	10.1	854	12.8
418–428	2.1	655–660	4.8	779–780	7.5		10.2	855–857	12.9
429–439	2.2	661–666	4.9	781–783	7.6	822	10.3	858–1400	12.9+
440–450	2.3	667–673	5.0	784–786	7.7		10.4		
451–460	2.4	674–679	5.1	787–788	7.8		10.5		
461–471	2.5	680–684	5.2	789–791	7.9	823	10.6		
472–481	2.6	685–690	5.3	792–793	8.0	824	10.7		

a. Extrapolated estimates were made for grade equivalents 0.0 to 0.9 based on the minimum expected Scaled Score for the Grade 1.0 grade equivalent.

b. Grade Equivalent scale scores for 11.0 to 12.9+ are based on the 2002 norms.

Table 58: Scaled Score to Percentile Ranks Conversion by Grade (at Month 10 [June] in the School Year)

	Grade									
PR	1	2	3	4	5	6	7	8	9	10
1	183	273	330	376	417	460	461	460	494	454
2	210	317	374	421	469	509	516	520	537	513
3	227	343	399	449	497	541	546	557	575	553
4	240	360	415	472	517	566	570	582	598	579
5	253	374	430	488	534	584	591	604	618	609
6	263	386	441	501	549	599	607	619	635	627
7	273	395	453	514	561	611	620	632	649	641
8	281	402	463	524	572	623	632	643	662	654
9	288	410	472	533	580	633	643	655	675	667
10	294	415	479	541	589	643	652	666	685	682
11	298	419	486	547	597	648	658	675	691	692
12	303	423	492	553	603	653	664	683	697	697
13	308	427	497	557	608	657	669	690	702	703
14	312	431	502	562	613	661	675	695	707	708
15	316	434	508	566	618	665	680	701	711	712
16	320	438	512	571	623	670	685	707	715	716
17	323	440	516	574	627	674	689	712	720	720
18	327	444	520	578	631	678	693	716	724	725
19	330	447	525	582	635	682	697	720	729	730
20	334	450	529	585	638	685	701	725	734	734
21	337	453	532	589	642	689	704	729	737	739
22	340	455	536	593	645	692	708	732	741	742
23	343	457	539	596	648	695	711	736	744	746
24	346	459	542	599	652	698	714	739	748	749
25	349	461	545	602	654	702	717	743	751	752
26	351	464	549	605	657	705	720	746	755	755
27	353	466	552	608	659	708	723	749	757	759

Table 58: Scaled Score to Percentile Ranks Conversion by Grade (at Month 10 [June] in the School Year) (Continued)

	Grade									
PR	1	2	3	4	5	6	7	8	9	10
28	356	468	554	611	663	711	726	752	761	763
29	358	470	557	613	666	714	730	756	764	767
30	360	472	559	616	669	716	733	759	768	771
31	362	474	562	618	672	718	736	763	772	773
32	365	476	565	621	674	721	738	766	775	776
33	367	478	567	623	677	724	742	769	777	780
34	369	480	570	626	680	726	745	771	781	784
35	372	482	572	628	682	728	748	775	783	787
36	374	485	575	630	685	731	751	778	786	790
37	376	487	577	632	687	733	754	781	789	793
38	378	489	579	635	689	736	757	784	791	796
39	380	491	581	637	692	739	761	786	794	799
40	383	493	583	639	694	742	764	788	796	802
41	386	495	585	642	696	745	768	791	798	804
42	388	497	588	645	698	747	771	794	802	806
43	391	499	590	647	701	750	774	797	804	809
44	393	501	593	649	704	753	777	800	807	811
45	395	504	595	652	706	755	781	803	809	813
46	397	506	597	654	708	758	784	806	812	816
47	399	508	599	656	710	761	787	809	814	820
48	401	510	601	658	713	764	789	813	816	822
49	403	511	603	660	715	767	791	815	820	825
50	406	513	605	663	717	769	794	819	822	828
51	409	515	607	665	719	772	796	822	825	831
52	412	518	609	668	721	775	799	825	827	834
53	414	520	611	670	723	777	802	827	829	836
54	416	522	613	672	725	780	804	830	832	839
55	418	525	615	674	727	782	807	833	835	842

Table 58: Scaled Score to Percentile Ranks Conversion by Grade (at Month 10 [June] in the School Year) (Continued)

	Grade									
PR	1	2	3	4	5	6	7	8	9	10
56	420	527	617	676	729	785	809	836	837	845
57	423	529	619	678	731	787	812	839	840	847
58	425	532	621	680	733	789	814	842	842	849
59	427	534	623	682	734	791	816	844	845	852
60	430	536	625	684	736	793	819	846	848	854
61	432	537	627	687	738	795	821	849	851	857
62	434	539	629	689	741	797	824	851	853	859
63	437	541	632	691	743	800	826	854	856	862
64	439	543	634	694	745	802	828	857	859	864
65	441	545	636	696	748	804	831	861	861	867
66	444	548	638	698	750	806	833	864	865	869
67	447	550	641	701	752	808	835	868	869	872
68	449	552	644	703	754	810	837	871	872	875
69	452	554	646	705	757	813	840	875	876	879
70	454	556	648	708	760	815	842	879	879	883
71	457	559	650	710	763	817	845	881	882	887
72	459	561	652	713	765	820	848	884	885	890
73	461	563	654	715	769	822	851	887	888	893
74	464	566	656	717	771	825	853	891	892	896
75	467	569	658	720	774	828	856	894	895	900
76	469	571	661	723	777	831	860	897	899	904
77	472	573	664	726	780	833	864	900	903	907
78	474	576	666	729	783	836	867	904	906	911
79	477	578	669	731	786	839	871	907	910	915
80	480	580	671	734	789	842	875	910	913	921
81	483	584	674	736	792	846	879	914	916	926
82	487	587	677	739	795	850	884	918	920	929
83	492	592	680	742	798	854	889	922	925	933

Table 58: Scaled Score to Percentile Ranks Conversion by Grade (at Month 10 [June] in the School Year) (Continued)

	Grade									
PR	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
84	496	596	683	745	802	859	893	927	929	936
85	500	600	686	749	805	864	897	931	933	942
86	504	604	690	752	809	869	902	935	938	947
87	509	608	693	756	813	874	908	941	943	952
88	514	612	696	759	817	879	912	947	949	961
89	517	616	699	763	822	885	918	953	957	967
90	523	621	703	768	827	891	922	960	967	975
91	528	626	708	774	832	897	927	966	972	982
92	533	632	714	781	839	904	932	974	980	989
93	538	638	720	788	846	910	938	982	986	1000
94	544	645	727	794	854	917	945	991	992	1008
95	551	652	733	802	863	927	953	1001	998	1018
96	561	660	742	812	873	937	964	1012	1009	1037
97	576	671	752	823	885	949	980	1027	1022	1059
98	597	685	766	837	900	966	999	1052	1043	1075
99	623	710	789	859	926	1001	1030	1084	1075	1120

Table 59: 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		

Table 60: Normal Curve Equivalent to Percentile Rank Conversions

NCE Range Low–High	PR	NCE Range Low–High	PR	NCE Range Low–High	PR	NCE Range Low–High	PR
1.0–4.0	1	36.1–36.7	26	50.3–50.7	51	64.6–65.1	76
4.1–8.5	2	36.8–37.3	27	50.8–51.2	52	65.2–65.8	77
8.6–11.7	3	37.4–38.0	28	51.3–51.8	53	65.9–66.5	78
11.8–14.1	4	38.1–38.6	29	51.9–52.3	54	66.6–67.3	79
14.2–16.2	5	38.7–39.2	30	52.4–52.8	55	67.4–68.0	80
16.3–18.0	6	39.3–39.8	31	52.9–53.4	56	68.1–68.6	81
18.1–19.6	7	39.9–40.4	32	53.5–53.9	57	68.7–69.6	82
19.7–21.0	8	40.5–40.9	33	54.0–54.4	58	69.7–70.4	83
21.1–22.3	9	41.0–41.5	34	54.5–55.0	59	70.5–71.3	84
22.4–23.5	10	41.6–42.1	35	55.1–55.5	60	71.4–72.2	85
23.6–24.6	11	42.2–42.7	36	55.6–56.1	61	72.3–73.1	86
24.7–25.7	12	42.8–43.2	37	56.2–56.6	62	73.2–74.1	87
25.8–26.7	13	43.3–43.8	38	56.7–57.2	63	74.2–75.2	88
26.8–27.6	14	43.9–44.3	39	57.3–57.8	64	75.3–76.3	89
27.7–28.5	15	44.4–44.9	40	57.9–58.3	65	76.4–77.5	90
28.6–29.4	16	45.0–45.4	41	58.4–58.9	66	77.6–78.8	91
29.5–30.2	17	45.5–45.9	42	59.0–59.5	67	78.9–80.2	92
30.3–31.0	18	46.0–46.5	43	59.6–60.1	68	80.3–81.7	93
31.1–31.8	19	46.6–47.0	44	60.2–60.7	69	81.8–83.5	94
31.9–32.6	20	47.1–47.5	45	60.8–61.3	70	83.6–85.5	95
32.7–33.3	21	47.6–48.1	46	61.4–61.9	71	85.6–88.0	96
33.4–34.0	22	48.2–48.6	47	62.0–62.5	72	88.1–91.0	97
34.1–34.7	23	48.7–49.1	48	62.6–63.1	73	91.1–95.4	98
34.8–35.4	24	49.2–49.7	49	63.2–63.8	74	95.5–99.0	99
35.5–36.0	25	49.8–50.2	50	63.9–64.5	75		

Appendix A: Objectives and STAR Math Items

Table 61: Numeration Concepts

NA1	Ones: Placing numerals in order
NA2	Ones: Using numerals to indicate quantity
NA3	Ones: Relate numerals and number words
NA4	Ones: Use ordinal numbers
N00	Ones: Locate numbers on a number line
N01	Tens: Place numerals (10–99) in order of value
N02	Tens: Associate numeral with group of objects
N03	Tens: Relate numeral and number word
N04	Tens: Identify one more/one less across decades
N05	Tens: Understand the concept of zero
N06	Hundreds: Place numerals in order of value
N07	Hundreds: Relate numeral and number word
N08	Hundreds: Identify place value of digits
N09	Hundreds: Write numerals in expanded form
N11	Thousands: Place numerals in order of value
N12	Thousands: Relate numeral and number word
N13	Thousands: Identify place value of digits
N14	Thousands: Write numerals in expanded form
N16	Ten thousands, hundred thousands, millions, billions: Place numerals in order of value
N17	Ten thousands, hundred thousands, millions, billions: Relate numeral and number word
N18	Ten thousands, hundred thousands, millions, billions: Identify place value of digits
N19	Ten thousands, hundred thousands, millions, billions: Write numerals in expanded form
N21	Fractions and decimals: Convert fraction to equivalent fraction
N22	Fractions and decimals: Convert fraction to decimal
N23	Fractions and decimals: Convert decimal to fraction

Table 61: Numeration Concepts (Continued)

N24	Fractions and decimals: Read word names for decimals to thousandths
N25	Fractions and decimals: Identify place value of digits in decimals
N26	Fractions and decimals: Identify position of decimals on number line
N27	Fractions and decimals: Identify position of fractions on number line
N28	Fractions and decimals: Convert improper fraction to mixed number
N29	Fractions and decimals: Round decimals to tenths, hundredths
N30	Fractions and decimals: Relate decimals to percents
N31	Advanced concepts: Determine square roots of perfect squares
N32	Advanced concepts: Give approximate square roots of a number
N33	Advanced concepts: Recognize meaning of nth root
N34	Advanced concepts: Recognize meaning of exponents (2–10)
N35	Advanced concepts: Recognize meaning of negative exponents
N36	Advanced concepts: Recognize meaning of fractional exponents
N37	Advanced concepts: Can use scientific notation
N38	Advanced concepts: Knows meaning of primes and composites
N39	Advanced concepts: Can determine greatest common factor
N40	Advanced concepts: Can determine least common multiple
N41	Advanced concepts: Recognizes use of negative numbers

Table 62: Computation Processes

C01	Addition of basic facts to 10
C02	Subtraction of basic facts to 10
C03	Addition of basic facts to 18
C04	Subtraction of basic facts to 18
C05	Addition of three single-digit addends
C06	Addition beyond basic facts, no regrouping (2d + 1d)
C07	Subtraction beyond basic facts, no regrouping (2d – 1d)
C08	Addition beyond basic facts with regrouping (2d + 1d, 2d + 2d)
C09	Subtraction beyond basic facts with regrouping (2d – 1d, 2d – 2d)

Table 62: Computation Processes (Continued)

C10	Addition beyond basic facts with double regrouping (3d + 2d, 3d + 3d)
C11	Subtraction beyond basic facts with double regrouping (3d – 2d, 3d – 3d)
C12	Multiplication basic facts
C13	Division basic facts
C14	Multiplication beyond basic facts, no regrouping (2d × 1d)
C15	Division beyond basic facts, no remainders (2d ÷ 1d)
C16	Multiplication with regrouping (2d × 1d, 2d × 2d)
C17	Division with remainders (2d ÷ 1d, 3d ÷ 1d)
C18	Addition of whole numbers: any difficulty
C19	Subtraction whole numbers: any difficulty
C21	Division of whole numbers: any difficulty
C22	Addition of fractions: like single-digit denominators
C23	Subtraction of fractions: like single-digit denominators
C24	Addition of fractions: unlike single-digit denominators
C25	Subtraction of fractions: unlike single-digit denominators
C26	Multiplication of fractions: single-digit denominators
C27	Division of fractions: single-digit denominators
C28	Addition of mixed numbers
C29	Subtraction of mixed numbers
C30	Multiplication of mixed numbers
C31	Division of mixed numbers
C33	Addition of decimals, place change (e.g. 2 + .45)
C35	Subtraction of decimals, place change (e.g. 5 – .4)
C36	Multiplication of decimals
C37	Division of decimals
C38	Percent A (10 is what % of 40?)
C39	Percent B (20% of 50 is what?)
C40	Percent C (30 is 50% of what?)
C41	Proportions
C42	Ratios

Table 63: Other Applications

Estimation	
E06	Estimation problems: Addition beyond basic facts, no regrouping (2d + 1d)
E07	Estimation problems: Subtraction beyond basic facts, no regrouping (2d – 1d)
E14	Estimation problems: Multiplication beyond basic facts, no regrouping (2d × 1d)
E15	Estimation problems: Division beyond basic facts, no remainders (2d ÷ 1d)
E18	Estimation problems: Addition of whole numbers, any difficulty
E19	Estimation problems: Subtraction of whole numbers, any difficulty
E20	Estimation problems: Multiplication of whole numbers, any difficulty
E21	Estimation problems: Division of whole numbers, any difficulty
E24	Estimation problems: Addition of fractions, unlike single-digit denominators
E25	Estimation problems: Subtraction of fractions, unlike single-digit denominators
E28	Estimation problems: Addition of mixed numbers
E29	Estimation problems: Subtraction of mixed numbers
E32	Estimation problems: Addition of decimals, no place change (e.g. 2.34 + 10.32)
E33	Estimation problems: Addition of decimals, place change (e.g. 2 + .45)
E34	Estimation problems: Subtraction of decimals, no place change (e.g. .53 –.42)
E35	Estimation problems: Subtraction of decimals, place change (e.g. 5 – .4)
E38	Estimation problems: Percent A (10 is what % of 40?)
E39	Estimation problems: Percent B (20% of 50 is what?)
E40	Estimation problems: Percent C (30 is 50% of what?)
Geometry	
GA1	Use basic terms to describe position
GA2	Identify common plane shapes
GA3	Identify common plane shapes when rotated
GA4	Compare common objects to basic shapes
GA5	Understand basic symmetry
GA6	Recognize elements of basic shapes
GA7	Identify common solid shapes
G00	Identify fraction parts of common plane shapes

Table 63: Other Applications (Continued)

G01	Identify numeric patterns
G02	Circle terms
G03	Perimeter: square
G04	Perimeter: rectangle
G05	Perimeter: triangle
G06	Area: square
G07	Area: rectangle
G08	Area: right triangle
G09	Area: circle
G10	Volume: rectangular prism
G12	Identify rays
G13	Identify line segments
G14	Identify parallel lines
G15	Identify intersecting lines
G16	Identify perpendicular lines
G17	Use properties of parallel lines
G18	Use properties of intersecting lines
G19	Use properties of perpendicular lines
G20	Vertical and supplementary angles
G21	Classify angles (obtuse, etc.)
G22	Using parts of a triangle
G23	Pythagorean theorem
Measurement	
MA1	Use simple vocabulary of measurement
MA2	Understand the value of penny, nickel, dime
MA3	Determine the value of quarter and dollar
MA4	Understand the value of groups of coins to $1.00
MA5	Tell time to the hour and half hour
MA6	Read a thermometer
MA7	Order days of the week

Table 63: Other Applications (Continued)

M00	Order months of the year	
M01	Customary measures: Inches, feet, yards	
M02	Customary measures: Estimating linear measures	
M03	Customary measures: Estimating volume measures	
M04	Customary measures: Pints, quarts, gallons	
M05	Metric prefixes	
M06	Metric: Customary conversions	
M07	Measures of angles	
M08	Estimating linear measure in metric units	
Data Analysis and Statistics		
SA1	Read tally charts	
S00	Read simple pictographs	
S01	Read table	
S02	Read bar graph	
S03	Read pie graph	
S04	Interpret table	
S05	Interpret bar graph	
S06	Interpret pie graph	
S07	Statistics: Mean	
S08	Statistics: Median	
S11	Probability: Simple	
S12	Probability: Joint	
Word Problems		
W03	Word problems: Addition of basic facts	
W04	Word problems: Subtraction of basic facts	
W06	Word problems: Addition beyond basic facts, no regrouping (2d + 1d)	
W08	Word problems: Addition beyond basic facts with regrouping (2d + 1d, 2d + 2d)	
W09	Word problems: Subtraction beyond basic facts with regrouping (2d – 1d, 2d – 2d)	
W12	Word problems: Multiplication of basic facts	
W13	Word problems: Division of basic facts	

Table 63: Other Applications (Continued)

W14	Word problems: Multiplication beyond basic facts, no regrouping (2d × 1d)
W15	Word problems: Division beyond basic facts, no remainders (2d ÷ 1d)
W16	Word problems: Multiplication with regrouping (2d × 1d, 2d × 2d)
W17	Word problems: Division with remainders (2d ÷ 1d, 3d ÷ 1d)
W18	Word problems: Addition of whole numbers, any difficulty
W19	Word problems: Subtraction of whole numbers, any difficulty
W20	Word problems: Multiplication of whole numbers, any difficulty
W21	Word problems: Division of whole numbers, any difficulty
W22	Word problems: Addition of fractions, like single-digit denominators
W23	Word problems: Subtraction of fractions, like single-digit denominators
W24	Word problems: Addition of fractions, unlike single-digit denominators
W25	Word problems: Subtraction of fractions, unlike single-digit denominators
W28	Word problems: Addition of mixed numbers
W29	Word problems: Subtraction of mixed numbers
W2S	Word problems: Two-step
W33	Word problems: Addition of decimals, place change (e.g. 2 + .45)
W35	Word problems: Subtraction of decimals, place change (e.g. 5 – .4)
W36	Word problems: Multiplication of decimals
W37	Word problems: Division decimals
W38	Word problems: Percent A (10 is what % of 40?)
W39	Word problems: Percent B (20% of 50 is what?)
W40	Word problems: Percent C (30 is 50% of what?)
W41	Word problems: Proportions
W42	Word problems: Ratios
WXI	Word problems: Extra information
Algebra	
A00	Can skip count by 2, 5, 10 in ascending order
A01	Simple number sentence
A02	Translate word problem to equation
A03	Linear equations: 1 unknown

Table 63: Other Applications (Continued)

A04	Linear equations: 2 unknowns
A05	Reciprocals of rational numbers
A06	Graph of linear equation (integers add, subtract)
A07	Linear inequalities: 1 unknown
A08	Linear inequalities: 2 unknown
A09	Graph linear inequalities
A10	Classify mono-, bi-, or trinomials
A11	Polynomials: Order polynomials
A12	Polynomials: Addition and subtraction
A13	Polynomials: Multiplication and division
A14	Solve system of 2 equations (2 unknowns)
A15	Quadratic equations: Solve using square root rule
A16	Quadratic equations: Solve by factoring
A17	Quadratic equations: Completing the square
A18	Factor common term from binomial expression
A19	Determine slope
A20	Determine intercept
A21	Sequences and series: Common differences in arithmetic sequences
A22	Sequences and series: Find specified term of arithmetic sequences
A25	Determine if functions are one to one (using graphs)
A26	Graph simple ellipses

Appendix B: Objectives and STAR Math Enterprise Items ENTERPRISE

STAR Math Enterprise is a skills-based assessment of math achievement in four domains and 54 skill sets for greater depth of assessment in grades 1–12. STAR Math Enterprise assesses skills in four standards-based math domains:

1. numbers and operations
2. algebra
3. geometry and measurement
4. data analysis, statistics and probability.

Within each domain, skills are organized into sets of closely related skills. The resulting hierarchical structure is domain, skill set, and skill. There are four math domains, 54 skill sets, and more than 550 skills.

Table 64: Numbers and Operations

Count with Objects and Numbers	
N56	Count objects to 20
NA1	Complete a sequence of numbers to 10
N42	Count on by ones from a number less than 100
N43	Count back by ones from a number less than 20
NA4	Answer a question involving an ordinal number up to “tenth”
N57	Identify a number to 20 represented by a point on a number line
N82	Locate a number to 20 on a number line
N58	Determine one more than or one less than a given number
N04	Determine one more than or one less than a given number across decades
N95	Determine ten more than or ten less than a given number
N59	Count by 2s to 50 starting from a multiple of 2
N96	Count by 5s or 10s to 100 starting from a multiple of 5 or 10, respectively
N02	Count objects grouped in tens and ones
N45	Complete a skip pattern starting from a multiple of 2, 5, or 10
NFY	Complete a skip pattern of 2 or 5 starting from any number
NFZ	Complete a skip pattern of 10 starting from any number
N46	Count on by 100s from any number

Table 64: Numbers and Operations (Continued)

Identify Odd and Even Numbers	
N97	Identify odd and even numbers less than 100
Relate Place and Value to a Whole Number	
N83	Determine the value of a digit in a 2-digit number
N74	Represent a 2-digit number as tens and ones
N98	Determine the 2-digit number represented as tens and ones
N03	Relate a whole number to the word form of the number to 100
N61	Compare whole numbers to 100 using words
N62	Order whole numbers to 100 in ascending order
N08	Identify the place of a digit in a 3-digit number
N84	Represent a 3-digit number as hundreds, tens, and ones
N64	Determine the 3-digit number represented as hundreds, tens, and ones
NAB	Recognize equivalent forms of a 3-digit number using hundreds, tens, and ones
N09	Represent a 3-digit whole number in expanded form
N07	Relate a 3-digit whole number to its word form
N76	Compare whole numbers to 1,000 using the symbols <, >, and =
N06	Order whole numbers to 1,000 in ascending or descending order
N48	Determine the value of a digit in a 4- or 5-digit whole number
N49	Determine which digit is in a specified place in a 4- or 5-digit whole number
NAE	Represent a 4-digit whole number as thousands, hundreds, tens, and ones
N86	Determine the 4-digit whole number represented in thousands, hundreds, tens, and ones
N14	Represent a 4-digit whole number in expanded form
N12	Relate a 4- or 5-digit whole number to its word form
N11	Order 4-digit whole numbers in ascending or descending order
N18	Determine the value of a digit in a 6-digit number
NAF	Determine the 4- or 5-digit whole number represented in expanded form
N19	Represent a 5-digit whole number in expanded form
N16	Order 4- to 6-digit whole numbers in ascending or descending order
N70	Round a 4- to 6-digit whole number to a specified place

Table 64: Numbers and Operations (Continued)

N17	Relate a 7- to 10-digit whole number to the word form of the number
N37	Convert a whole number greater than 10 to scientific notation
Add and Subtract Whole Numbers without Regrouping	
N99	Determine equivalent forms of a number, up to 10
A38	Determine the missing portion in a partially screened (hidden) collection of up to 10 objects
C43	Know basic addition facts to 10 plus 10
C44	Know basic subtraction facts to 20 minus 10
W03	WP: Use basic addition facts to solve problems
W04	WP: Use basic subtraction facts to solve problems
C06	Add a 2-digit number and a 1-digit number without regrouping
W06	WP: Add a 2-digit number and a 1-digit number without regrouping
C67	Add two 2-digit numbers without regrouping
C07	Subtract a 1-digit number from a 2-digit number without regrouping
N05	Add or subtract zero to or from any number less than 100
C87	Subtract a 2-digit number from a 2-digit number without regrouping
WXP	WP: Subtract a 1-digit number from a 2-digit number without regrouping
WXQ	WP: Add two 2-digit numbers without regrouping
WXR	WP: Subtract a 2-digit number from a 2-digit number without regrouping
WXU	WP: Determine a basic addition-fact number sentence for a given situation
WXV	WP: Determine a basic subtraction-fact number sentence for a given situation
WXW	WP: Add two 3-digit numbers without regrouping
WXY	WP: Subtract a 3-digit number from a 3-digit number without regrouping
E41	Estimate a sum or difference of 2- to 4-digit whole numbers using any method
W7B	WP: Estimate a sum or difference of two 3- or 4-digit whole numbers using any method
Add and Subtract Whole Numbers with Regrouping	
C05	Add three 1-digit numbers
C88	Determine a number pair that totals 100
C08	Add a 2-digit number and a 1- or 2-digit number with regrouping
W08	WP: Add a 2-digit number and a 1- or 2-digit number with regrouping

Table 64: Numbers and Operations (Continued)

C47	Add 2- and 3-digit numbers with no more than one regrouping
C69	Add two 3-digit numbers with one regrouping
C09	Subtract a 1- or 2-digit number from a 2-digit number with one regrouping
W09	WP: Subtract a 1- or 2-digit number from a 2-digit number with one regrouping
C70	Subtract a 1- or 2-digit number from a 3-digit number with one regrouping
C71	Subtract a 3-digit number from a 3-digit number with one regrouping
C49	Add 3- and 4-digit whole numbers with regrouping
C18	Add four 1- to 4-digit whole numbers
C11	Subtract a 2- or 3-digit number from a 3-digit number with two regroupings
C50	Subtract 3- and 4-digit whole numbers with regrouping
C19	Subtract two 2- to 6-digit whole numbers
W18	WP: Add 3- and 4-digit whole numbers with regrouping
W19	WP: Subtract 3- and 4-digit whole numbers with regrouping
Multiply Whole Numbers	
CE0	Know multiplication tables for 2, 5, and 10
C72	Use a multiplication sentence to represent an area or an array model
C91	Know basic multiplication facts to 10 × 10
W65	WP: Multiply using basic facts to 10 × 10
C14	Multiply a 2-digit whole number by a 1-digit whole number with no regrouping
W14	WP: Multiply a 2-digit whole number by a 1-digit whole number without regrouping
E14	Estimate the product of a 2-digit number and a 1-digit number
C52	Multiply a 1- or 2-digit whole number by a multiple of 10, 100, or 1,000
C53	Apply the distributive property to multiply a multi-digit number by a 1-digit number
C54	Multiply a 3- or 4-digit whole number by a 1-digit whole number
C16	Multiply a 2-digit whole number by a 1- or 2-digit whole number with regrouping
W16	WP: Multiply a 2-digit whole number by a 1- or 2-digit whole number
C74	Multiply a 2-digit whole number by a 2-digit whole number
W46	WP: Multiply a multi-digit whole number by a 1-digit whole number
E20	Estimate the product of whole numbers using any method
W8F	WP: Estimate a product of two whole numbers using any method

Table 64: Numbers and Operations (Continued)

W20	WP: Multiply whole numbers
W51	WP: Solve a multi-step problem involving whole numbers
Divide Whole Numbers without a Remainder in the Quotient	
W53	WP: Divide objects into equal groups by sharing
C73	Know basic division facts to 100 ÷ 10
W66	WP: Divide using basic facts to 100 ÷ 10
C15	Divide a 2-digit whole number by a 1-digit whole number with no remainder in the quotient
W15	WP: Divide a 2-digit whole number by a 1-digit whole number with no remainder in the quotient
W2S	WP: Solve a 2-step whole number problem using more than one operation
W90	WP: Divide a 3-digit whole number by a 1-digit whole number with no remainder in the quotient
E15	Estimate the quotient of a 2-digit whole number divided by a 1-digit whole number with no remainder in the quotient
C21	Divide whole numbers with no remainder in the quotient
W21	WP: Divide whole numbers with no remainder in the quotient
E21	Estimate a quotient using any method
W58	WP: Estimate a quotient using any method
Divide Whole Numbers with a Remainder in the Quotient	
C17	Divide a 2- or 3-digit whole number by a 1-digit whole number with a remainder in the quotient
W7C	WP: Divide a 3-digit whole number by a 1-digit whole number with a remainder in the quotient
W17	WP: Divide a 2- or 3- digit whole number by a 1-digit whole number with a remainder in the quotient
C55	Divide a multi-digit whole number by a 2-digit whole number, with a remainder and at least one zero in the quotient
C56	Divide a multi-digit whole number by a 2-digit whole number and express the quotient as a mixed number
W49	WP: Solve a 2-step problem involving whole numbers
W57	WP: Divide a whole number and interpret the remainder

Table 64: Numbers and Operations (Continued)

Identify, Compare, and Order Fractions	
N87	Determine a pictorial model of a fraction of a whole
N67	Determine a pictorial model of a fraction of a set of objects
N77	Identify a fraction represented by a point on a number line
N68	Locate a fraction on a number line
N78	Compare fractions using models
N88	Order fractions using models
N69	Identify equivalent fractions using models
N27	Locate a mixed number on a number line
N21	Identify a fraction equivalent to a given fraction
N91	Compare fractions with unlike denominators
NB3	Order fractions with unlike denominators in ascending or descending order
Add and Subtract Fractions with Like Denominators	
C22	Add fractions with like 1-digit denominators
W22	WP: Add fractions with like denominators no greater than 10 and simplify the sum
C23	Subtract fractions with like 1-digit denominators
W23	WP: Subtract fractions with like denominators no greater than 10
WCE	WP: Subtract fractions with like denominators no greater than 10 and simplify the difference
WXZ	WP: Add fractions with like denominators and simplify the sum
WX2	WP: Subtract fractions with like denominators and simplify the difference
WX3	WP: Add mixed numbers with like denominators and simplify the sum
WX4	WP: Subtract mixed numbers with like denominators and simplify the difference
Find Prime Factors, Common Factors, and Common Multiples	
N38	Identify the prime factors of a 2-digit number
N39	Determine the greatest common factor of two whole numbers
N40	Determine the least common multiple of two whole numbers
Add and Subtract Fractions with Unlike Denominators	
C57	Add fractions with unlike denominators that have factors in common and simplify the sum
C24	Add fractions with unlike 1-digit denominators

Table 64: Numbers and Operations (Continued)

W24	WP: Add fractions with unlike 1-digit denominators
C76	Add fractions with unlike denominators that have no factors in common
C28	Add mixed numbers with unlike denominators
E28	Estimate the sum of mixed numbers
W28	WP: Add mixed numbers with unlike denominators
C77	Subtract fractions with unlike denominators that have factors in common and simplify the difference
C25	Subtract fractions with unlike 1-digit denominators
W25	WP: Subtract fractions with unlike 1-digit denominators
C78	Subtract fractions with unlike denominators that have no factors in common
C29	Subtract mixed numbers with unlike denominators
W29	WP: Subtract mixed numbers with unlike denominators
E24	Estimate the sum of fractions with unlike 1-digit denominators
E25	Estimate the difference between fractions with unlike 1-digit denominators
E29	Estimate the difference between mixed numbers with unlike denominators
Convert between an Improper Fraction and a Mixed Number	
N72	Convert a mixed number to an improper fraction
N28	Convert an improper fraction to a mixed number
Relate a Decimal to a Fraction	
NB2	Determine the fraction equivalent to a decimal number model
N22	Convert a fraction or mixed number in hundredths or thousandths to a decimal number
NB1	Determine the decimal number equivalent to a fraction model
N23	Convert a decimal number in hundredths or thousandths to a fraction
N81	Compare numbers in decimal and fractional forms
Relate Place and Value to a Decimal Number	
NB9	Determine the decimal number from a pictorial model of tenths or hundredths
N71	Identify a pictorial model of tenths or hundredths of a decimal number
NBA	Identify a decimal number to tenths represented by a point on a number line
N51	Locate a decimal number to tenths on a number line
N50	Read a decimal number through the hundredths place

Table 64: Numbers and Operations (Continued)

N79	Compare decimal numbers through the hundredths place
N89	Order decimal numbers through the hundredths place
N24	Relate a decimal number through ten-thousandths to its word form
N29	Round a decimal number to a specified place through hundredths
N25	Identify the place of a digit in a decimal number through hundredths
N80	Compare decimal numbers of differing places to thousandths
NB5	Order decimal numbers of differing places to thousandths in ascending or descending order
N54	Represent a decimal number in expanded form using powers of ten
N55	Determine the decimal number represented in expanded form using powers of ten
N26	Estimate a decimal number from its position on a number line
N92	Order numbers in decimal and fractional forms
NB7	Convert a number less than 1 to scientific notation
NB8	Convert a number less than 1 from scientific notation to standard form
Add or Subtract Decimal Numbers	
W54	WP: Determine the amount of change from whole dollar amounts
C51	Determine money amounts that total $10
C33	Determine the sum of a whole number and a decimal number to hundredths
W33	WP: Determine the sum of a decimal number and a whole number
E33	Estimate the sum of a whole number and a decimal number
E32	Estimate the sum of two decimal numbers
C98	Add two decimal numbers of differing places to thousandths
W96	WP: Estimate the sum or difference of two decimal numbers through thousandths using any method
E45	Estimate the sum of two decimal numbers through thousandths and less than 1 by rounding to a specified place
C79	Add decimal numbers and whole numbers
C93	Subtract two decimal numbers of differing places to thousandths
W94	WP: Add or subtract decimal numbers through thousandths
W95	WP: Add or subtract a decimal number through thousandths and a whole number
E34	Estimate the difference of two decimal numbers

Table 64: Numbers and Operations (Continued)

E44	Estimate the difference of two decimal numbers through thousandths and less than 1 by rounding to a specified place
C35	Subtract a decimal number from a whole number
W35	WP: Subtract a decimal number from a whole number
E35	Estimate the difference of a whole number and a decimal number
Divide a Whole Number Resulting in a Decimal Quotient	
C58	Divide a whole number by a 1-digit whole number resulting in a decimal quotient through thousandths
C59	Divide a whole number by a 2-digit whole number resulting in a decimal quotient through thousandths
W50	WP: Divide a whole number by a 1- or 2-digit whole number resulting in a decimal quotient
Multiply and Divide with Fractions	
C26	Multiply a fraction by a fraction
C80	Multiply a mixed number by a whole number
C61	Multiply a mixed number by a fraction
ABF	Determine the reciprocal of a positive whole number, a proper fraction, or an improper fraction
C27	Divide a fraction by a fraction
C82	Divide a whole number by a fraction resulting in a fractional quotient
W71	WP: Multiply or divide two mixed numbers or a mixed number and a fraction
C81	Divide a fraction by a whole number resulting in a fractional quotient
W59	WP: Multiply or divide a fraction by a fraction
W99	WP: Solve a 2-step problem involving fractions
WA9	WP: Solve a multi-step problem involving fractions or mixed numbers
C30	Multiply mixed numbers
C31	Divide mixed numbers
Multiply and Divide with Decimals	
C94	Multiply a decimal number through thousandths by 10, 100, or 1,000
C9F	Multiply a decimal number through thousandths by a whole number
W80	WP: Multiply a decimal number through thousandths by a whole number
W60	WP: Estimate the product of two decimals

Table 64: Numbers and Operations (Continued)

C36	Multiply two decimal numbers
W36	WP: Multiply two decimal numbers
C83	Multiply decimal numbers less than one in hundredths or thousandths
CA0	Multiply decimal numbers greater than one where the product has 2 or 3 decimal places
C99	Divide a decimal number by 10, 100, or 1,000
C84	Divide a decimal number through thousandths by a 1- or 2-digit whole number where the quotient has 2–5 decimal places
W9B	WP: Divide a decimal number through thousandths by a 1- or 2-digit whole number
C9A	Divide a 1- to 3-digit whole number by a decimal number to tenths where the quotient is a whole number
W9C	WP: Divide a whole number by a decimal number through thousandths, rounded quotient if needed
C85	Divide a 1- to 3-digit whole number by a decimal number to tenths where the quotient is a decimal number to thousandths
C9B	Divide a 2- or 3-digit whole number by a decimal number to hundredths or thousandths, rounded quotient if needed
C37	Divide decimal numbers
W37	WP: Divide a whole number by a decimal number
C86	Divide a decimal number by a decimal number through thousandths, rounded quotient if needed
W81	WP: Divide a decimal through thousandths by a decimal through thousandths, rounded quotient if needed
W9D	WP: Estimate the quotient of two decimals
W9E	WP: Solve a 2-step problem involving decimals
W86	WP: Solve a multi-step problem involving decimal numbers
Relate a Decimal Number to a Percent	
NFT	Convert a decimal number to a percentage
N30	Convert a percentage to its decimal equivalent
N0W	Convert a decimal number in thousandths to a percentage
Solve a Proportion, Rate, or Ratio	
WA2	WP: Use a unit rate, with a whole number or whole cent value, to solve a problem
W82	WP: Determine a unit rate with a whole number value

Table 64: Numbers and Operations (Continued)

C42	Determine if ratios are equivalent
W42	WP: Determine if ratios are equivalent
WA0	WP: Determine a part given a ratio and the whole where the whole is less than 50
WA1	WP: Determine the whole given a ratio and a part where the whole is less than 50
C41	Solve a proportion involving whole numbers
W41	WP: Solve a proportion
C38	Determine the percent a whole number is of another whole number
E38	Estimate the percent a whole number is of another whole number
C39	Determine a given percent of a number
E39	Estimate a given percent of a number
C40	Determine a whole number given a part and a percent
E40	Estimate a whole number given a part and a percent
WAC	WP: Determine a unit rate
WAD	WP: Use a unit rate to solve a problem
W88	WP: Determine a part, given part to whole ratio and the whole, where the whole is greater than 50
WAA	WP: Determine a part, given part to part ratio and the whole, where the whole is greater than 50
W89	WP: Determine a part, given part to whole ratio and a part, where the whole is greater than 50
WAB	WP: Determine a part, given part to part ratio and a part, where the whole is greater than 50
W8A	WP: Determine the whole, given part to whole ratio and a part, where the whole is greater than 50
W73	WP: Determine the whole, given part to part ratio and a part, where the whole is greater than 50
Evaluate a Numerical Expression	
N93	Evaluate a numerical expression of four or more operations, with parentheses, using order of operations
N34	Evaluate or represent an expression of a whole number raised to a whole number power
NB6	Evaluate an integer raised to a whole number power
N94	Evaluate a numerical expression involving integer exponents and/or integer bases

Table 64: Numbers and Operations (Continued)

N35	Evaluate a whole number raised to a negative power
A49	Evaluate a numerical expression involving one or more exponents and multiple forms of rational numbers
AA1	Simplify a monomial numerical expression involving the square root of a whole number
N33	Evaluate the nth root of a whole number
N36	Evaluate a whole number raised to a fractional power
Perform Operations with Integers	
C62	Add integers
C63	Subtract integers
C64	WP: Add and subtract using integers
C65	Multiply integers
C66	Divide integers
W87	WP: Multiply or divide integers
Determine a Square Root	
N31	Evaluate the positive square root of a perfect square
NBB	Determine the square root of a perfect-square fraction or decimal
N32	Determine an approximate square root of a number
NFV	Determine both square roots of a perfect square
NBC	Determine the two closest integers to a given square root
NBD	Approximate the location of a square root on a number line
Solve a Problem Involving Percents	
WA6	WP: Determine the percent of decrease applied to a number
WA7	WP: Determine the percent of increase applied to a number
W84	WP: Determine the result of applying a percent of decrease to a value
WA8	WP: Determine the result of applying a percent of increase to a value
W85	WP: Answer a question involving a fraction and a percent
C97	Determine a percent of a number given a percent that is not a whole percent
W8B	WP: Determine a given percent of a number
W8C	WP: Determine the percent one number is of another number

Table 64: Numbers and Operations (Continued)

W8D	WP: Determine a number given a part and a decimal percentage or a percentage more than 100%
WB1	WP: Estimate a given percent of a number
W38	WP: Determine the percent a whole number is of another whole number, with a result less than 100%
W39	WP: Determine a percent of a whole number using percents less than 100
W40	WP: Determine a whole number given a part and a percent
C9C	Determine the percent one number is of another number
C9D	Determine a number given a part and a decimal percentage or a percentage more than 100%

Table 65: Algebra

Relate a Rule to a Pattern	
A39	Determine the rule for an addition or subtraction number pattern
A29	Extend a number pattern involving addition
A95	Extend a number pattern involving subtraction
A40	Identify a missing figure in a growing pictorial or nonnumeric pattern
A44	Generate a table of paired numbers based on a rule
A31	Identify a missing term in a multiplication or a division number pattern
AA4	Determine a rule that relates two variables
A32	Determine the variable expression with one operation for a table of paired numbers
W97	WP: Determine the variable expression with one operation for a table of paired numbers
W7E	WP: Generate a table of paired numbers based on a variable expression with one operation
A21	Determine the common difference in an arithmetic sequence
A22	Find a specified term in an arithmetic sequence
Determine the Operation Given a Situation	
A30	WP: Determine the operation needed for a given situation
W67	WP: Determine a multiplication or division sentence for a given situation
C90	Use a division sentence to represent objects divided into equal groups

Table 65: Algebra (Continued)

Graph on a Coordinate Plane	
GFS	Determine the ordered pair of a point in the first quadrant
GFV	Determine the ordered pair of a point in any quadrant
AAC	Use a table to represent the values from a first-quadrant graph
A48	Determine the graph of a 1-operation linear function
AA7	Determine the graph of a 2-operation linear function
AA8	Determine the slope of a line given its graph or a graph of a line with a given slope
AA0	Determine the graph of a line using given information
A52	Determine the graph of a linear equation given in slope-intercept, point-slope, or standard form
A91	Determine the graph of a given quadratic function
W79	WP: Answer a question using the graph of a quadratic function
A08	Determine the graph of a 2-variable linear inequality
A25	Relate a graph to an equation of a parabola
A26	Relate a graph of an ellipse to its equation
Evaluate an Algebraic Expression or Function	
A33	Evaluate a 2-variable expression, with two or three operations, using whole number substitution
W72	WP: Evaluate a 1- or 2-variable expression or formula using whole numbers
A36	Evaluate a 2-variable expression, with two or three operations, using integer substitution
A50	Evaluate a function written in function notation for a given value
Solve a Linear Equation	
A28	Determine a missing addend in a basic addition-fact number sentence
WXS	WP: Determine a missing addend in a basic addition-fact number sentence
A81	Determine a missing subtrahend in a basic subtraction-fact number sentence
WXT	WP: Determine a missing subtrahend in a basic subtraction-fact number sentence
A01	Determine a missing addend in a number sentence involving 2-digit numbers
AF5	Determine the reciprocal of a negative rational number
A45	Solve a 1-step equation involving whole numbers
A47	Solve a 1-step linear equation involving integers

Table 65: Algebra (Continued)

A43	Solve a 2-step linear equation involving integers
A37	Solve a proportion involving decimals
A98	Solve a 1-step equation involving rational numbers
A99	Solve a 2-step equation involving rational numbers
W75	WP: Solve a problem involving a 1-variable, 2-step equation
A51	Solve a 1-variable linear equation with the variable on both sides
AAB	Rewrite an equation to solve for a specified variable
A04	Determine a solution to a 2-variable linear equation
Determine a Linear Equation	
WA3	WP: Use a 2-variable equation to represent a situation involving a direct proportion
W83	WP: Use a 2-variable linear equation to represent a situation
A42	Use a 2-variable equation to construct an input-output table
A46	Use a 2-variable equation to represent a relationship expressed in a table
A02	Use a 1-variable, 1-step equation to represent a verbal statement
WAF	WP: Use a 1-variable 1-step equation to represent a situation
AA5	Determine the table of values that represents a linear equation with rational coefficients in two variables
AA6	Determine a linear equation in two variables that represents a table of values
W8E	WP: Use a 1-variable equation with rational coefficients to represent a situation involving two operations
WB2	WP: Use a 2-variable equation with rational coefficients to represent a situation
A9C	Determine the slope-intercept form or the standard form of a linear equation
A53	Determine an equation of a line given the slope and y-intercept of the line
A06	Determine an equation for a line given a graph
A83	Determine an equation for a line given the slope of the line and a point on the line that is not the y-intercept
A84	Determine an equation of a line given two points on the line
Identify Characteristics of a Linear Equation or Function	
AA9	Determine the x- or y-intercept of a line given its graph
W76	WP: Interpret the meaning of the slope of a graphed line
WB3	WP: Interpret the meaning of the y-intercept of a graphed line

Table 65: Algebra (Continued)

A20	Determine the x- or y-intercept of a line given an equation
A19	Determine the slope of a line given two points on the line or the graph of the line
A9A	WP: Determine a reasonable domain or range for a function in a given situation
A9E	Determine the slope of a line given an equation of the line
Solve a System of Linear Equations	
A14	Solve a system of linear equations in two variables using any method
Determine a System of Linear Equations	
W74	WP: Determine a system of linear equations that represents a given situation
Simplify an Algebraic Expression	
A61	Simplify an algebraic expression by combining like terms
A97	Multiply two monomial algebraic expressions
A13	Multiply two binomials
A18	Factor a common term from a binomial expression
A87	Apply the product of powers property to a monomial algebraic expression
A88	Apply the power of a power property to a monomial algebraic expression
A89	Apply the power of a product property to a monomial algebraic expression
A8A	Apply the quotient of powers property to monomial algebraic expressions
A8B	Apply the power of a quotient property to monomial algebraic expressions
A12	Add or subtract polynomial expressions
A8E	Multiply two binomials of the form (ax +/– b)(cx +/– d)
A8F	Factor the GCF from a polynomial expression
A90	Factor trinomials that result in factors of the form (ax +/– b)(cx +/– d)
AA2	Simplify a monomial algebraic radical expression
A55	Simplify a rational expression involving polynomial terms
A56	Multiply rational expressions
A57	Divide a polynomial expression by a monomial
A58	Add or subtract two rational expressions with unlike polynomial denominators
Solve a Linear Inequality	
A07	Determine the solution set of a 1-variable linear inequality
AAA	Solve a 2-step linear inequality in one variable

Table 65: Algebra (Continued)

WB4	WP: Solve a problem involving a 2-step linear inequality in one variable
A62	Determine the graph of the solutions to a 2-step linear inequality in one variable
A9B	Solve a 1-variable linear inequality with the variable on both sides
Solve a Nonlinear Equation	
A85	Solve a 1-variable absolute value inequality
AA3	Solve a radical equation that leads to a linear equation
A93	Solve a quadratic equation using the quadratic formula
A16	Solve a quadratic equation by factoring
A15	Solve a quadratic equation using the square root rule
A54	Solve a radical equation that leads to a quadratic equation
A59	Solve a rational equation involving terms with monomial denominators
A60	Solve a rational equation involving terms with polynomial denominators
Graph a 1-Variable Inequality	
A09	Relate a 1-variable inequality to its graph

Table 66: Geometry and Measurements

Relate Money to Symbols, Words, and Amounts	
MA2	Identify a coin or the value of a coin
C89	Determine cent amounts that total a dollar
MA4	Determine the value of groups of coins to $1.00
N75	Translate between a dollar sign and a cent sign
NAC	Convert money amounts in words to amounts in symbols
Use the Vocabulary of Geometry and Measurement	
MA1	Compare objects using the vocabulary of measurement
GA6	Identify a shape with given attributes
GA2	Identify a circle, a triangle, a square, or a rectangle
GA1	Use basic terms to describe position
G37	Determine the common attributes in a set of geometric shapes
GA7	Identify a common solid shape
GA5	Identify a line of symmetry
G14	Identify parallel lines
G15	Identify intersecting line segments

Table 66: Geometry and Measurements (Continued)

GFZ	Classify a right angle or a straight angle given a picture
G21	Classify an obtuse angle or an acute angle given a picture
G30	Classify an angle given its measure
G13	Identify line segments
G19	Identify perpendicular or parallel lines when given a transversal
G16	Identify perpendicular lines
G12	Identify rays
Determine a Missing Figure in a Pattern	
G01	Identify a missing figure in a geometric pattern
A96	Identify a missing figure in a repeating pictorial or nonnumeric pattern
Determine a Measurement	
MA9	Measure length in inches
M09	Measure length in centimeters
MAA	Read a thermometer in degrees Fahrenheit or Celsius
M02	Estimate the height or length of a common object in customary units
M08	Estimate the height of a common object in metric units
M01	Convert between inches, feet, and yards
M04	Convert between customary units of capacity
M05	Convert within metric units of mass, length, and capacity
M07	Identify an angle given its measure
M06	Determine the approximate value of a unit converted between customary and metric measures
M18	WP: Determine a measure of length, weight or mass, or capacity or volume using proportional relationships
M11	Convert a rate from one unit to another with a change in one unit
M12	Convert a rate from one unit to another with a change in both units
G20	Determine the measure of a vertical angle or a supplementary angle
G18	Identify angle relationships formed by intersecting lines
G17	Identify angle relationships formed by parallel lines cut by a transversal
Tell Time	
MA5	Tell time to the hour and half hour
M15	Tell time to the quarter hour
M16	Tell time to 5-minute intervals
M10	Tell time to the minute

Table 66: Geometry and Measurements (Continued)

Calculate Elapsed Time	
M17	Calculate elapsed time exceeding an hour with regrouping
W68	WP: Calculate elapsed time exceeding an hour with regrouping hours
Solve a Problem Involving the Perimeter of a Shape	
G05	Determine the perimeter of a triangle
GAB	Determine the perimeter of a rectangle given a picture showing length and width
G03	Determine the perimeter of a square
G04	WP: Determine the perimeter of a rectangle
GAC	Determine the missing side length of a rectangle given a side length and the perimeter
WA4	WP: Determine the perimeter or the area of a complex shape
G26	Solve a problem involving the circumference of a circle
Solve a Problem Involving the Area of a Shape	
GAD	Determine the area of a polygon on a grid
G06	Determine the area of a square
G07	Determine the area of a rectangle given the length and width
W56	WP: Determine the area of a rectangle
GAF	Determine the missing side length of a rectangle given a side length and the area
W98	WP: Determine the area of a square or rectangle
G08	Determine the area of a right triangle
G24	Use a formula to determine the area of a triangle
W69	WP: Determine the area of a triangle
W70	WP: Determine a missing dimension given the area and another dimension
G25	Determine the area of a complex shape
GE5	Determine the area of a right triangle or a rectangle given the coordinates of the vertices of the figure
GGS	Determine the area of a quadrilateral
GGT	Determine a length given the area of a quadrilateral
G09	Determine the area of a circle
G33	Solve a problem given the area of a circle
GGU	Determine the area of a sector of a circle
GGV	Determine the length of the radius or the diameter of a circle given the area of a sector
GGW	WP: Determine a length or an area involving a sector of a circle
GGX	Determine the measure of an arc or an angle given the area of a sector of a circle

Table 66: Geometry and Measurements (Continued)

Identify Congruence and Similarity of Geometric Shapes	
GA4	Compare common objects to basic shapes
GA3	Identify figures that are the same size and shape
GA8	Determine lines of symmetry
GB0	Determine the result of a flip, a turn, or a slide
GE7	Identify a triangle congruence postulate that justifies a congruence statement
GFF	Identify congruent triangles using triangle congruence postulates or theorems
GF7	Identify a triangle similarity postulate that justifies a similarity statement
GF8	Identify similar triangles using triangle similarity postulates or theorems
Solve a Problem Involving the Surface Area or Volume of a Solid	
G10	Determine the volume of a rectangular prism
W7F	WP: Determine the volume of a rectangular prism
G31	Determine the surface area of a rectangular prism
G32	WP: Find the surface area of a rectangular prism
G34	Determine the volume of a rectangular or a triangular prism
W61	WP: Solve a problem involving the volume of a geometric solid
W62	WP: Determine the surface area of a geometric solid
GGY	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
GH0	Solve a problem involving the volume of a right pyramid or a right cone
GH1	Determine the surface area of a sphere
GH2	Determine the volume of a sphere or hemisphere
Determine a Missing Measure or Dimension of a Shape	
G02	Relate the radius to the diameter in a circle
G27	Determine a missing dimension given two similar shapes
WB0	WP: Solve a problem involving similar shapes
G22	Determine a missing angle measure in a triangle
G23	Use the Pythagorean theorem to determine a length
WB5	WP: Use the Pythagorean theorem to find a length or a distance
GFG	Solve a problem involving the distance formula
GE4	Determine the midpoint of a line segment given the coordinates of the endpoints
GE6	Determine the measure of an angle formed by parallel lines and one or more transversals
GFH	Solve a problem using inequalities in a triangle

Table 66: Geometry and Measurements (Continued)

GG4	WP: Determine a length or an angle measure using triangle relationships
GF6	Determine the measure of an angle or the sum of the angles in a polygon
GG6	WP: Solve a problem using the properties of angles and/or sides of polygons
GF9	Determine a length using parallel lines and proportional parts
GGE	WP: Determine a length using similarity
GFJ	Determine a length in a complex figure using the Pythagorean theorem
GFA	Determine a length using the properties of a 45-45-90 degree triangle or a 30-60-90 degree triangle
GGP	Determine the measure of an arc or a central angle using the relationship between the arc and the central angle
GFB	Solve a problem involving the length of an arc
GFC	Determine the length of a line segment, the measure of an angle, or the measure of an arc using a tangent to a circle
GFD	Determine a length using a line segment tangent to a circle and the radius that intersects the tangent
GFE	Determine the measure of an arc or an angle using the relationship between an inscribed angle and its intercepted arc

Table 67: Data Analysis, Statistics, and Probability

Read or Answer a Question about Charts, Tables, or Graphs	
SA1	Read a tally chart
SD7	Read a 2-category tally chart
SD9	Answer a question using information from a 2-category tally chart
S00	Read a simple pictograph
S18	Answer a question using information from a pictograph (1 symbol = more than 1 object)
S01	Read a table
S04	Answer a question using information from a table
S02	Read a bar graph
S05	Answer a question using information from a bar graph
S19	Answer a question using information from a bar graph with a y-axis scale by 2s
SDC	Read a line plot
SDD	Answer a question using information from a line plot
S03	Read a circle graph

Table 67: Data Analysis, Statistics, and Probability (Continued)

S06	Answer a question using information from a circle graph
SA2	Read a line graph
S13	Answer a question using information from a line graph
S21	Read a double-bar graph
S22	Answer a question using information from a double-bar graph
S24	Answer a question using information from a histogram
SE6	Answer a question using information from a scatter plot
S23	Answer a question using information from a circle graph using percentage calculations
Use a Chart, Table, or Graph to Represent Data	
SD8	Use a 2-category tally chart to represent groups of objects (1 symbol = 1 object)
S17	Use a pictograph to represent data (1 symbol = more than 1 object)
S26	Use a bar graph with a y-axis scale by 2s to represent data
SD1	Use a line plot to represent data
S20	Use a line graph to represent data
SA3	Use a double-bar graph to represent data
S15	Use a circle graph to represent percentage data
S16	Use a histogram to represent data
SD5	Use a scatter plot to organize data
Determine a Measure of Central Tendency	
S07	Determine the mean of a set of whole number data
S14	Determine the median of an odd number of data values
SD3	Determine the median of an even number of data values
S08	Determine the median of a set of data given a frequency table
Use a Proportion to Make an Estimate	
S25	Use a proportion to make an estimate, related to a population, based on a sample
Determine the Probability of One or More Events	
S11	Determine the probability of a single event
S12	Determine the probability of independent events

Appendix C: Algebra Readiness Skills

The math concepts and skills learned in elementary through middle school provide the foundation for studying high-school-level algebra. The STAR Math Student Instructional Planning Report provides an Algebra Readiness Indicator to help teachers identify student progress through these foundational skills.

Research has identified the progression of skills needed for algebra readiness. The following tables list the Accelerated Math Second Edition skills associated with these algebra readiness skills. These lists can help teachers identify the grade-level skills a student may need to practice in order to achieve expected grade-level progress.

Table 68: Accelerated Math Second Edition Skills, Grade 3

Objective Number	Objective Name
13	Add 3- and 4-digit whole numbers with regrouping
14	Add three 2- to 3-digit whole numbers
15	Subtract 3- and 4-digit whole numbers with regrouping
16	WP: Add or subtract 3- and 4-digit whole numbers with regrouping
33	Use a multiplication sentence to represent an area or an array model
34	Use a division sentence to represent objects divided into equal groups
35	Know basic multiplication facts to 10 × 10
36	Know basic multiplication facts for 11 and 12
37	Know basic division facts to 100 ÷ 10
38	Know basic division facts for 11 and 12
39	WP: Multiply using basic facts to 10 × 10
40	WP: Divide using basic facts to 100 ÷ 10
41	Complete a multiplication and division fact family
43	Multiply a 2-digit whole number by a 1-digit number
44	Determine a pictorial model of a fraction of a whole
45	Determine a pictorial model of a fraction of a set of objects
46	Identify a fraction represented by a point on a number line
47	Locate a fraction on a number line
48	Compare fractions using models
50	Identify equivalent fractions using models

Table 68: Accelerated Math Second Edition Skills, Grade 3 (Continued)

Objective Number	Objective Name
51	Compare fractions with like denominators
52	Compare fractions with like numerators
53	WP: Compare equal unit fractions of different-sized wholes
58	Determine the missing multiplicand in a number sentence involving basic facts
59	Determine the missing dividend or divisor in a number sentence involving basic facts
62	WP: Determine the operation needed for a given situation
63	WP: Determine a multiplication or division sentence for a given situation
65	Determine a rule for a table of related number pairs
66	WP: Find the missing number in a table of paired values
103	Determine a location using map coordinates
104	Determine the map coordinates for a location

Table 69: Accelerated Math Second Edition Skills, Grade 4

Objective Number	Objective Name
21	Apply the distributive property to the multiplication of a 2-digit number by a 1- or 2-digit number
22	Apply the distributive property to multiply a multi-digit number by a 1-digit number
25	Multiply a 3-digit whole number by a 2-digit whole number
29	WP: Multiply a 3-digit whole number by a 2-digit whole number
33	Divide a multi-digit whole number by 10 or 100 with no remainder
37	Divide a 3-digit whole number by a 1-digit whole number with a remainder in the quotient
41	WP: Divide a 3-digit whole number by a 1-digit whole number with a remainder in the quotient
43	Identify a mixed number represented by a model
44	Identify a mixed number represented by a point on a number line
45	Locate a mixed number on a number line
46	WP: Use a mixed number to represent an amount in a sharing situation

Table 69: Accelerated Math Second Edition Skills, Grade 4 (Continued)

Objective Number	Objective Name
47	Identify an improper fraction represented by a model of a mixed number
48	Identify an improper fraction represented by a point on a number line
49	Locate an improper fraction on a number line
50	Simplify a fraction
51	Determine a set of equivalent fractions
52	Compare fractions on a number line
56	Add fractions with like denominators no greater than 10 and simplify the sum
57	WP: Add fractions with like denominators no greater than 10 and simplify the sum
60	Subtract fractions with like denominators no greater than 10 and simplify the difference
61	WP: Subtract fractions with like denominators no greater than 10 and simplify the difference
64	Determine the decimal number from a pictorial model of tenths or hundredths
65	Identify a pictorial model of tenths or hundredths of a decimal number
66	Identify a decimal number to tenths represented by a point on a number line
67	Locate a decimal number to tenths on a number line
68	Determine the decimal number equivalent to a fraction with a denominator of 10 or 100
69	Determine a fraction equivalent to a decimal, using a denominator of 10 or 100
70	Determine the decimal number equivalent to a fraction model
71	Determine the fraction equivalent to a decimal number model
72	Compare decimal numbers through the hundredths place
75	Add two decimal numbers through hundredths
76	Subtract two decimal numbers through hundredths
79	WP: Add or subtract decimal numbers of the same place through hundredths
84	Evaluate a numeric expression involving two operations

Table 69: Accelerated Math Second Edition Skills, Grade 4 (Continued)

Objective Number	Objective Name
85	Solve a 1-step addition or subtraction equation using a model
86	Identify a missing term in a multiplication or a division number pattern
90	Generate a table of paired numbers based on a rule
91	Determine a rule that relates two variables
92	Extend a number pattern in a table of related pairs
110	Determine the perimeter of a rectangle given a picture showing length and width
112	WP: Determine the perimeter of a square or rectangle
113	Determine the missing side length of a rectangle given a side length and the perimeter
116	Determine the area of a rectangle given the length and width
117	WP: Determine the area of a rectangle
118	Determine the missing side length of a rectangle given a side length and the area

Table 70: Accelerated Math Second Edition Skills, Grade 5

Objective Number	Objective Name
4	Determine the prime factorization of a number to 50
5	Determine the common factors for two whole numbers to 50
6	Determine the greatest common factor of two whole numbers to 50
7	Determine the multiple(s) of a number
8	Determine common multiples for two whole numbers
9	Determine the least common multiple of two whole numbers
13	Divide a multi-digit whole number by multiples of 100 or 1,000
21	Divide a multi-digit whole number by a 2-digit whole number, with a remainder and at least one zero in the quotient
22	Divide a multi-digit whole number by a 2-digit whole number and express the quotient as a mixed number
24	WP: Divide a whole number and interpret the remainder
25	WP: Solve a 2-step problem involving whole numbers

Table 70: Accelerated Math Second Edition Skills, Grade 5 (Continued)

Objective Number	Objective Name
31	Determine equivalent fractions not in simplest form
32	Determine the simplest form of a fraction
33	Compare fractions with unlike denominators
38	Add fractions with unlike denominators that have factors in common and simplify the sum
39	Add fractions with unlike denominators that have no factors in common
43	Subtract fractions with unlike denominators that have factors in common and simplify the difference
44	Subtract fractions with unlike denominators that have no factors in common
46	WP: Add or subtract fractions with unlike denominators that have no factors in common
47	Convert a mixed number to an improper fraction
48	Convert an improper fraction to a mixed number
50	Add mixed numbers with unlike denominators and simplify the sum
52	Subtract mixed numbers with unlike denominators and simplify the difference
54	WP: Add or subtract mixed numbers with unlike denominators that have no factors in common
60	Multiply a whole number by a unit fraction
62	Multiply a proper fraction by a whole number
64	Divide a whole number by a unit fraction
65	Divide a unit fraction by a whole number
67	Divide a whole number by a fraction, with a whole number quotient
68	WP: Multiply or divide a whole number by a unit fraction
73	Compare decimal numbers of differing places to thousandths
75	Add two decimal numbers of differing places to thousandths
77	Add decimal numbers and whole numbers
78	Subtract two decimal numbers of differing places to thousandths
79	Subtract a decimal number from a whole number or a whole number from a decimal number

Table 70: Accelerated Math Second Edition Skills, Grade 5 (Continued)

Objective Number	Objective Name
80	WP: Add or subtract decimal numbers through thousandths
81	WP: Add or subtract a decimal number through thousandths and a whole number
86	Multiply a decimal number through thousandths by 10, 100, or 1,000
87	WP: Multiply a decimal through thousandths by 10, 100, or 1,000
91	Convert a fraction with a denominator that is a factor of 10, 100, or 1,000 to decimal notation
94	Relate an equivalent fraction and percent given a grid
95	Relate an equivalent decimal and percent given a grid
96	Evaluate a numerical expression involving three operations, with no parentheses, using order of operations
97	Evaluate a numerical expression involving three operations, with parentheses, using order of operations
98	Use a variable expression with one operation to represent a verbal expression
100	WP: Use a variable expression with one operation to represent a situation
101	Evaluate a 1-variable expression, involving one operation, using whole number substitution
102	Evaluate a 2-variable expression, involving one operation, using whole number substitution
103	WP: Evaluate a 1-variable expression with one operation using a whole number value
104	WP: Evaluate a 2-variable expression with one operation using whole number values
107	Generate a table of paired numbers based on a variable expression with two operations
108	Determine the variable expression with one operation for a table of paired numbers
109	WP: Generate a table of paired numbers based on a variable expression with one operation
110	WP: Determine the variable expression with one operation for a table of paired numbers
111	Use a first quadrant graph to represent the values from a table generated in context

Table 70: Accelerated Math Second Edition Skills, Grade 5 (Continued)

Objective Number	Objective Name
126	Use a formula to determine the area of a triangle
127	Determine the area of a complex figure divided into basic shapes
128	Use a formula to determine the area of a parallelogram
129	WP: Determine the area of a triangle
130	WP: Determine the area of a square or rectangle
131	WP: Determine a missing dimension given the area and another dimension
134	Determine the volume of a rectangular prism
135	WP: Determine the volume of a rectangular prism
139	Determine the surface area of a rectangular prism
140	WP: Find the surface area of a rectangular prism
150	Determine the location of an ordered pair in the first quadrant
151	Determine the ordered pair of a point in the first quadrant

Table 71: Accelerated Math Second Edition Skills, Grade 6

Objective Number	Objective Name
2	Determine the greatest common factor of three numbers to 100
3	Determine the least common multiple of three numbers
4	WP: Determine the least common multiple of two or more numbers
7	Determine the square of a whole number to 15
8	Determine the cube of a whole number to 15
10	Divide a whole number by a 2-digit whole number resulting in a decimal quotient through thousandths
11	WP: Divide a whole number by a 1- or 2-digit whole number resulting in a decimal quotient
12	WP: Solve a multi-step problem involving whole numbers
17	Add mixed numbers with unlike denominators or a mixed number and a fraction with unlike denominators and simplify the sum
19	Subtract mixed numbers with unlike denominators or a mixed number and a fraction and simplify the difference

Table 71: Accelerated Math Second Edition Skills, Grade 6 (Continued)

Objective Number	Objective Name
21	WP: Add or subtract mixed numbers with unlike denominators or a mixed number and a fraction with unlike denominators and simplify the sum or difference
22	Multiply a fraction by a fraction
25	Multiply a mixed number by a mixed number
26	Determine the reciprocal of a whole number, a proper fraction, or an improper fraction
27	Determine the reciprocal of a mixed number
28	Divide a fraction by a whole number resulting in a fractional quotient
29	Divide a fraction by a fraction
30	Divide a whole number by a fraction resulting in a fractional quotient
31	Divide a mixed number by a fraction
32	Divide a mixed number by a mixed number
33	WP: Multiply or divide a fraction by a fraction
34	WP: Multiply or divide two mixed numbers or a mixed number and a fraction
35	WP: Solve a 2-step problem involving fractions
43	Multiply a decimal number through thousandths by a whole number
44	WP: Multiply a decimal number through thousandths by a whole number
48	Multiply decimal numbers less than one in hundredths or thousandths
50	Multiply decimal numbers greater than one where the product has 2 or 3 decimal places
51	WP: Multiply two decimal numbers to thousandths
53	Divide a decimal number by 10, 100, or 1,000
55	Divide a decimal number through thousandths by a 1- or 2-digit whole number where the quotient has 2–5 decimal places
57	Divide a whole number or a decimal number by 0.1, 0.01, or 0.001
61	Divide a 1- to 3-digit whole number by a decimal number to tenths where the quotient is a decimal number to thousandths
62	Divide a 2- or 3-digit whole number by a decimal number to hundredths or thousandths, rounded quotient if needed

Table 71: Accelerated Math Second Edition Skills, Grade 6 (Continued)

Objective Number	Objective Name
63	Divide a decimal number by a decimal number through thousandths, rounded quotient if needed
64	WP: Divide a whole number by a decimal number through thousandths, rounded quotient if needed
65	WP: Divide a decimal through thousandths by a decimal through thousandths, rounded quotient if needed
67	WP: Solve a 2-step problem involving decimals
68	Convert a mixed number to a decimal number
69	Convert a decimal number to a mixed number
72	Convert a decimal number to a percentage
73	Convert a percentage to a decimal number
80	WP: Calculate the percent of a whole number where the answer is a whole number
81	WP: Determine a ratio using whole numbers less than 50
82	Determine if ratios, using whole numbers less than 50, are equivalent
83	WP: Determine a part given a ratio and the whole where the whole is less than 50
84	WP: Determine a part given a ratio and another part where the whole is less than 50
85	WP: Determine the whole given a ratio and a part where the whole is less than 50
86	WP: Determine a unit rate with a whole number value
87	WP: Use a unit rate, with a whole number or whole cent value, to solve a problem
92	Determine which property of addition or multiplication justifies a step in the simplification of an expression
96	WP: Use a 2-variable equation to represent a situation involving a direct proportion
97	WP: Use a 2-variable linear equation to represent a situation
98	Evaluate a 1-variable expression, with two or three operations, using whole number substitution
99	Evaluate a 2-variable expression, with two or three operations, using whole number substitution

Table 71: Accelerated Math Second Edition Skills, Grade 6 (Continued)

Objective Number	Objective Name
101	Solve a 1-step equation involving whole numbers
102	Solve a proportion
104	Use a 2-variable equation to construct an input-output table
105	Use a 2-variable equation to represent a relationship expressed in a table
106	Use a first quadrant graph to represent the values in an input-output table
107	Use a graph to determine the entries in an input-output table
125	Determine the circumference of a circle using 3.14 for π
126	WP: Determine the circumference of a circle
127	Determine the volume of a prism with a right triangle base

Table 72: Accelerated Math Second Edition Skills, Grade 7

Objective Number	Objective Name
1	Determine the exponential notation that represents a repeated multiplication
2	Determine the repeated multiplication that is represented by a number raised to a power
4	Evaluate the positive square root of a perfect square
5	Evaluate a whole number power of a whole number
7	Determine the prime factorization of a number using exponents
10	Evaluate an expression containing the fraction bar as the division sign
11	Evaluate a numerical expression, with parentheses and exponents, using order of operations
18	Determine a percent of a whole number using less than 100%
19	Determine a percent of a whole number using more than 100%
20	Determine the percent a whole number is of another whole number, with a result less than 100%
21	Determine a whole number given a part and a percentage less than 100%
22	WP: Determine a percent of a whole number using less than 100%
23	WP: Determine the percent a whole number is of another whole number, with a result less than 100%

Table 72: Accelerated Math Second Edition Skills, Grade 7 (Continued)

Objective Number	Objective Name
24	WP: Determine a whole number given a part and a percentage
25	WP: Determine the percent of decrease applied to a number
26	WP: Determine the percent of increase applied to a number
27	WP: Determine the result of applying a percent of decrease to a value
28	WP: Determine the result of applying a percent of increase to a value
31	WP: Solve a multi-step problem involving decimal numbers
32	WP: Solve a multi-step problem involving fractions or mixed numbers
37	Add integers
39	Subtract integers
40	WP: Add and subtract using integers
41	Multiply integers
42	Divide integers
43	WP: Multiply or divide integers
44	WP: Determine the ratio of two whole numbers, at least one of which is larger than 50
45	Determine ratios equivalent to a given ratio of two whole numbers, at least one of which is larger than 50
46	WP: Determine a part, given part to whole ratio and the whole, where the whole is greater than 50
47	WP: Determine a part, given part to part ratio and the whole, where the whole is greater than 50
48	WP: Determine a part, given part to whole ratio and a part, where the whole is greater than 50
49	WP: Determine a part, given part to part ratio and a part, where the whole is greater than 50
50	WP: Determine the whole, given part to whole ratio and a part, where the whole is greater than 50
51	WP: Determine the whole, given part to part ratio and a part, where the whole is greater than 50
52	WP: Determine a unit rate
53	WP: Use a unit rate to solve a problem

Table 72: Accelerated Math Second Edition Skills, Grade 7 (Continued)

Objective Number	Objective Name
57	Identify a positive or negative rational number represented by a point on a number line
58	Locate a positive or negative rational number on a number line
59	Compare rational numbers (positive and negative)
61	Evaluate a rational expression involving variables with two or more terms in the numerator or denominator
63	Evaluate a 2-variable expression, with two or three operations, using integer substitution
64	Evaluate an algebraic expression involving whole number exponents
70	Solve a proportion involving decimals
71	WP: Solve a proportion
73	Solve a 1-step linear equation involving integers
75	WP: Use a 1-variable 1-step equation to represent a situation
76	Determine the graph of an inequality on a number line
77	Determine some solutions to a 1-variable linear inequality
85	Solve a problem involving the circumference of a circle
86	Determine the area of a trapezoid
89	Determine the area of a circle using 3.14 for π
91	WP: Determine the area of a circle using 3.14 for π
92	Solve a problem given the area of a circle
94	Determine the volume of a cylinder
95	WP: Determine the volume of a cylinder
96	WP: Solve a problem involving the volume of a geometric solid
98	Determine the surface area of a triangular prism
99	Determine the surface area of a cylinder
100	WP: Determine the surface area of a geometric solid
104	Determine a missing dimension given two similar shapes
105	WP: Solve a problem involving similar shapes
136	Use a proportion to make an estimate, related to a population, based on a sample

Table 73: Accelerated Math Second Edition Skills, Grade 8

Objective Number	Objective Name
6	Add or subtract signed fractions or mixed numbers
7	Multiply or divide signed fractions or mixed numbers
8	Add or subtract signed decimals
9	Multiply or divide signed decimals
10	Evaluate a numerical expression involving nested parentheses
13	Determine both square roots of a perfect square
15	Approximate the location of a square root on a number line
19	Compare rational numbers and/or irrational numbers in various forms
21	Determine a percent of a number given a percent that is not a whole percent
22	Determine the percent one number is of another number
23	Determine a number given a part and a decimal percentage or a percentage more than 100%
24	WP: Determine a given percent of a number
25	WP: Determine the percent one number is of another number
26	WP: Determine a number given a part and a decimal percentage or a percentage more than 100%
60	Determine the volume of a pyramid or a cone
61	WP: Determine the volume of a pyramid or a cone
62	Determine the surface area of a pyramid or a cone
63	WP: Determine the surface area of a pyramid or a cone
64	Solve a problem involving the surface area or the volume of a pyramid or a cone

References

Allington, R., & McGill-Franzen, A. (2003). Use students' summer-setback months to raise minority achievement. *Education Digest, 69*(3), 19–24.

Alonzo, A. C., & Gearhart, M. (2006). Considering learning progressions from a classroom assessment perspective. *Measurement, 14*(1 & 2), 99–104.

Bandeira de Mello, V., Blankenship, C., & McLaughlin, D. H. (2009). *Mapping state proficiency standards onto NAEP Scales: 2005–2007* (NCES 2010-456). Washington, DC: National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education.

Betebenner, D. W. (2009). Norm- and criterion-referenced student growth. *Educational Measurement: Issues and Practice, 28*(4), 42–51.

Betebenner, D. W. (2010). New directions for student growth models. Retrieved from the National Center for the Improvement of Educational Assessment website: http://www.ksde.org/LinkClick.aspx?fileticket=UssiNoSZks8%3D&tabid=4421&mid=10564

Betebenner, D. W., & Iwaarden, A. V. (2011a). SGP: An R package for the calculation and visualization of student growth percentiles [Computer Software manual]. (R package version 0.4-0.0 available at http://cran.r-project.org/web/packages/SGP/)

Betebenner, D. W. (2011b). A technical overview of the student growth percentile methodology: Student growth percentiles and percentile growth trajectories/projections. The National Center for the Improvement of Educational Assessment. Retrieved from http://www.nj.gov/education/njsmart/performance/SGP_Technical_Overview.pdf

Bock, R. D., Thissen, D., & Zimowski, M. F. (1997). IRT estimation of domain scores. *Journal of Educational Measurement, 34,* 197–211.

Bracey, G. (2002). Summer loss: The phenomenon no one wants to deal with. *Phi Delta Kappan, 84*(1), 12–13.

Corcoran, T., Mosher, F. A., & Rogat, A. (2009). *Learning progressions in science: An evidence-based approach to reform of teaching.* (CPRE Research Report # RR-63). New York: Consortium for Policy Research.

Cronin, J., Kingsbury, G. G., Dahlin, M., & Bowe, B. (2007). *Alternate methodologies for estimating state standards on a widely used computer adaptive test.* Paper presented at the American Educational Research Association, Chicago, IL.

Deno, S. (2003). Developments in curriculum-based measurement. *Journal of Special Education, 37*(3), 184–192.

Fuchs, D., & Fuchs, L. S. (2006). Introduction to Response to Intervention: What, why, and how valid is it? *Reading Research Quarterly, 41*(1), 93–99.

Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis.* Orlando, FL: Academic Press.

Heritage, M. (2008). *Learning progressions: Supporting instruction and formative assessment.* Washington, DC: Chief Council of State School Officers.

Kolen, M., & Brennan, R. (2004). *Test equating, scaling, and linking* (2nd ed.). New York: Springer.

Leahy, S., & Wiliam, D. (2011). *Devising learning progressions assessment.* Paper presented at the Annual Meeting of the American Educational Research Association, New Orleans, LA.

McLaughlin, D. H., & Bandeira de Mello, V. (2002). *Comparison of state elementary school mathematics achievement standards using NAEP 2000.* Paper presented at the annual meeting of the American Educational Research Association, New Orleans, LA.

McLaughlin, D. H., & Bandeira de Mello, V. (2003). *Comparing state reading and math performance standards using NAEP.* Paper presented at the National Conference on Large-Scale Assessment, San Antonio, TX.

McLaughlin, D. H., Bandeira de Mello, V., Blankenship, C., Chaney, K., Esra, P., Hikawa, H., et al. (2008). *Comparison between NAEP and state reading assessment results: 2003* (NCES 2008-474). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics.

Market Data Retrieval. (2001). A D&B Company: Shelton, CT.

Monaghan, W. (2006). *The facts about subscores* (ETS R&D Connections No. 4). Princeton, NJ: Educational Testing Services. Available online: http://www.ets.org/Media/Research/pdf/RD_Connections4.pdf.

National Council of Teachers of Mathematics. (2006). *Curriculum focal points for prekindergarten through grade 8 mathematics.* Reston, VA.

National Mathematics Advisory Panel. (2008). *Foundations for success: The final report of the National Mathematics Advisory Panel.* Washington, DC: U.S. Department of Education. Available online: http://www.ed.gov/about/bdscomm/list/mathpanel/report/final-report.pdf.

Neter, J., Kutner, M., Nachtsheim, C., & Wasserman, W. (1996). *Applied linear statistical models* (4th ed.). New York: WCB McGraw-Hill.

Pepe M. S., Janes, H., Longton, G. Leisenring, W., & Newcomb, P. (2004). Limitations of the odds ratio in gauging the performance of a diagnostic, prognostic, or screening marker. *American Journal of Epidemiology, 159,* 882–890.

Popham, W. J. (2007). The lowdown on learning progressions. *Educational Leadership, 64*(7), 83–84.

Skoupski, W.P., & Carvajal, J. (2010). A comparison of approaches for improving the reliability of objective scores. *Educational and Psychological Measurement, 70*(3), 357–375.

Urry, V. W. (1975). *The effects of guessing on parameters of item discriminatory power.* TN 75-2. Washington DC: Personnel Research and Development Center, US Civil Service Commission, May.

VanDerHeyden, A., & Burns, M. (2008). Examination of the Utility of Various Measures of Mathematics Proficiency. *Assessment for Effective Intervention, 33*(4), 215–224.

Zhou, X. H., Obuchowski, N. A., & Obushcowski, D. M. (2002). *Statistical methods in diagnostic medicine.* Wiley & Sons: New York.

Index

G

I

K

L

M

N

O

P

R

S

T

V

W

About Renaissance Learning

Renaissance Learning, Inc. is a leading provider of technology-based school improvement and student assessment programs for K12 schools. Renaissance Learning's tools provide daily formative assessment and periodic progress-monitoring technology to enhance core curriculum, support differentiated instruction, and personalize practice in reading, writing and math. Renaissance Learning products help educators make the practice component of their existing curriculum more effective by providing tools to personalize practice and easily manage the daily activities for students of all levels. As a result, teachers using Renaissance Learning products accelerate learning, get more satisfaction from teaching, and help students achieve higher test scores on state and national tests.

P.O. Box 8036 | Wisconsin Rapids, WI 54495-8036 | (800) 338-4204 | www.renlearn.com

Renaissance Learning™
Advanced Technology for Data-Driven Schools

L0319.0912
R54119