THE RESEARCH FOUNDATION FOR

STAR Assessments™

The Science of STAR

RENAISSANCE
LEARNING
Accelerating learning for all

STAR Early Literacy™, **STAR Math™**, and **STAR Reading™** are highly rated for progress monitoring by the National Center on Intensive Intervention.

National Center on Response to Intervention
www.rti4success.org

STAR Early Literacy™ is highly rated for screening and progress monitoring by the National Center on Response to Intervention.

STAR Reading™ and **STAR Math™** received the highest possible ratings for screening and progress monitoring from the National Center on Response to Intervention, with perfect scores in all categories.

Contents

Figures

Tables

Dear Educator,

Renaissance Learning is the world's leading provider of computer-based assessment technology, with products in use worldwide in grades pre-K–12. Renaissance Learning tools have a research base unmatched by makers of other educational products and have met the highest review standards set by reputable organizations such as the National Center on Intensive Intervention, the National Center on Response to Intervention, National Center on Student Progress Monitoring, the National Dropout Prevention Center, the Promising Practices Network, and the What Works Clearinghouse.
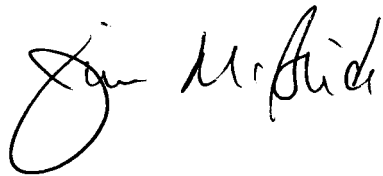
All Renaissance Learning tools are designed to accomplish our mission—"accelerating learning for all." A key educational principle supporting this mission is the notion that "the initial step in accelerating learning is to measure its occurrence." Our assessments—STAR Early Literacy Enterprise, STAR Reading Enterprise, and STAR Math Enterprise—do just that.

There is a reason approximately 18,000 schools worldwide use at least one STAR Enterprise assessment. They quickly gain favor with educators because of their ease of use, quick administration times, and ability to provide teachers with highly valid and reliable data upon completion of each test. The computer-based STAR assessment system is a multipurpose tool. STAR is used for screening and progress monitoring, and also includes resources that target instruction for all kinds of learners. Students who are most at risk can be identified quickly. No time is wasted in diagnosing their needs, allowing intervention to begin immediately.

Read on to learn more about STAR Enterprise assessments. I'm confident you'll see rather quickly why teachers using STAR Enterprise accelerate learning, get more satisfaction from teaching, and help their students achieve higher scores on state and national tests. The stakes are high. We must help all students in all schools be prepared for college or careers by the time they graduate from high school.

For additional information, full technical manuals are available for each STAR assessment by contacting Renaissance Learning at research@renlearn.com

Sincerely,

James R. McBride, Ph.D.
Vice President & Chief Psychometrician
Renaissance Learning, Inc.

James R. McBride, Ph.D., is vice president and chief psychometrician for Renaissance Learning. He was a leader of the pioneering work related to computerized adaptive testing (CAT) conducted by the Department of Defense. McBride has been instrumental in the practical application of item response theory (IRT) and since 1976 has conducted test development and personnel research for a variety of organizations. At Renaissance Learning, he has contributed to the psychometric research and development of STAR Math, STAR Reading, and STAR Early Literacy. McBride is co-editor of a leading book on the development of CAT and has authored numerous journal articles, professional papers, book chapters, and technical reports.

Introduction

STAR Enterprise assessments are designed to help teachers assess students quickly, accurately, and efficiently. STAR provides teachers with reliable and valid data instantly so that they can target instruction, monitor progress, provide students with the most appropriate instructional materials, and intervene with at-risk students. Administrators use real-time data from STAR to make decisions about curriculum, assessment, and instruction at the classroom, school, and district levels.

Three STAR Enterprise assessments measure student achievement in four areas:

- STAR Early Literacy Enterprise assesses early literacy and early numeracy skills (grades pre-K–3)
- STAR Reading Enterprise assesses reading skills (grades K–12)
- STAR Math Enterprise assesses math skills (grades K–12)

All STAR Enterprise assessments include skills-based test items, the Core Progress learning progressions for instructional planning, and in-depth reports. Operating on the Renaissance Place hosted platform, STAR Enterprise is a comprehensive assessment system for data-driven schools. The assessments provide accurate data in a short amount of time by combining computer-adaptive technology with a specialized psychometric test design that utilizes item response theory (IRT).

Students take STAR Enterprise assessments on individual computers or iPads. The software delivers multiple-choice items one by one, and a student selects answers with a mouse, keyboard, or touchscreen. After an assessment is completed, the software calculates the student's score. Teachers and administrators then select reports to provide results for an individual student, class, grade, school, or district.

STAR Assessments have been favorably reviewed as reliable, valid, and efficient by various independent groups, including the National Center on Intensive Intervention, the National Center on Response to Intervention, and the National Center on Student Progress Monitoring. STAR also has a significant research base as shown in Table 1.

Table 1: Research Support for STAR Assessments™

Assessment	Total Research Publications	Independent Research Publications
STAR Early Literacy	21	14
STAR Reading	76	22
STAR Math	65	21

STAR Assessments™ Overview

STAR Early Literacy Enterprise™ Assessment

The importance of assessing skills early in a child's schooling cannot be overstated. Research supports successful early intervention as the best single predictor for future academic success, particularly in the critical areas of reading and language acquisition.

Students are expected to develop a variety of early literacy as they progress from pre-kindergarten through third grade on their way to becoming readers. This progression reflects both the home literacy environment and educational interventions. The development of these skills, however, is not continuously upward. Students typically learn a skill, forget it, and then relearn it. Many well-established tests assess a student at a particular point in time. STAR Early Literacy Enterprise is designed to repeatedly assess a child's status at different stages throughout this important growth period.

STAR Early Literacy Enterprise measures early literacy and early numeracy skills throughout the early primary grades (pre-K–3).[1] Information from the assessment enables teachers to intervene immediately at the beginning of a student's formal learning process. This is particularly critical for students who enter school already lacking in experiences or the foundational skills necessary for early literacy and early numeracy development to take root.

STAR Early Literacy Enterprise is a standards-based test that measures student performance in key early literacy and early numeracy skills, providing valuable information regarding the acquisition of ability along a continuum of expectations. Table 2 breaks down the STAR Early Literacy Enterprise item bank by overall size, number of items administered per testing event, and average administration time.

Table 2: Summary of STAR Early Literacy Enterprise™ Item Bank Size and Administration Details

	STAR Early Literacy Enterprise
Item Bank Size	More than 2,500 items
Items Administered per testing event	27 items
Average Administration Time	About 10 minutes

For teachers, STAR Early Literacy Enterprise provides a simple way to monitor progress based on the specific needs of each student. It is especially helpful in identifying students who may be at risk for later reading failure. Data from the assessment is used for goal setting and outcome assessment as well as for planning instruction and intervention. A student's scaled score from STAR Early Literacy is also mapped to the empirically validated Core Progress learning progression. This score represents an entry point onto Core Progress, and using this tool, teachers can clearly see the skills students have likely mastered and the ones they are ready to develop next (for more about Core Progress, see pp. 10 and 31). Although STAR Early Literacy Enterprise is designed for students in grades pre-K–3, it can be used with older students, such as struggling readers, nonreaders, special education students, or English learners.

---

[1] STAR Early Literacy Enterprise is specifically designed for students who do not yet read. Students who have established a 100-sight-word vocabulary, or have reached the Probable Reader stage of literacy development in STAR Early Literacy Enterprise, typically are ready to take a STAR Reading Enterprise assessment.

STAR Early Literacy Enterprise is distinguished from other assessments of early literacy in three ways:

1. It is computer-administered, using graphics, audio instructions, and automatic dictation of instructions and test questions, so that most children can take the test without teacher assistance.
2. It is computer adaptive, which means the content and difficulty level of each test administration is tailored to each student's performance.
3. It is brief, administering 27 items (including five early numeracy items) in about 10 minutes. Despite its brevity, the assessment correlates highly with a wide range of more time-intensive standardized measures of early literacy, reading, and other readiness skills. Figure 1 shows sample assessment items.

Figure 1: STAR Early Literacy Enterprise™ Sample Assessment Items

Early Literacy Item
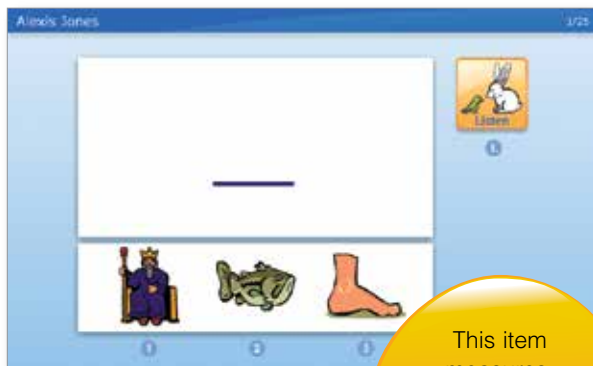
This item measures:
Sound-Symbol Correspondence: Consonants

Early Numeracy Item

$$4 - 1 = 3$$

| 1 | 2 | 3 |

"There are four cars in a parking lot. One car leaves. [...] number for how many cars are left."
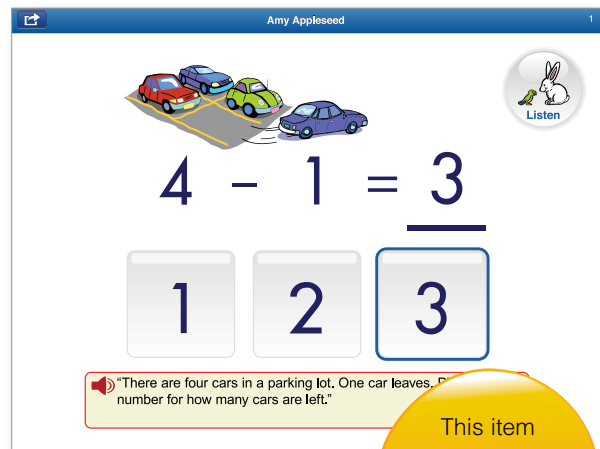
This item measures:
Composing and Decomposing

STAR Reading Enterprise™ Assessment

STAR Reading Enterprise is a challenging, interactive, and brief (about 15 minutes) assessment, consisting of 34 questions per test, that evaluates a breadth of reading skills appropriate for grades K–12.[2] The assessment's repeatability and flexibility in administration provide specific advantages for everyone responsible for the education of students:

- Teachers use results from STAR Reading Enterprise to facilitate individualized instruction and identify students who most need remediation or enrichment.
- Principals access assessment information through browser-based management and regular, accurate reports on performance at the individual, class, building, and district level.
- Administrators and assessment specialists apply reliable and timely information on reading growth at each school and districtwide, which serves as a valid basis for comparing data across schools, grades, and special student populations.

---

[2] Although STAR Reading Enterprise is normed for grades 1–12, kindergarten students may take the assessment with teacher discretion. Students with a 100-sight-word vocabulary, or who have reached the Probable Reader stage of literacy development in STAR Early Literacy Enterprise, are typically ready to take the assessment.

STAR Reading Enterprise is a standards-based test that measures student performance in key reading skills, providing valuable information regarding the acquisition of reading ability along a continuum of literary expectations. Table 3 breaks down the STAR Reading Enterprise item bank by overall size, number and types of items administered per testing event, and average administration time.

Table 3: Summary of STAR Reading Enterprise™ Item Bank Size and Administration Details

	STAR Reading Enterprise
Item Bank Size	More than 5,000
Items Administered per testing event	34 items
Average Administration Time	About 15 minutes

Renaissance Learning has conducted extensive research and consulted heavily with reading and assessment experts to arrive at the skills most appropriate for assessing reading development. Several publications have been studied, including the 2010 Common Core State Standards; the Reading Framework for the 2009 National Assessment of Educational Progress; the National Council of Teachers of English (2006) *Principles of Adolescent Literacy Reform* policy brief; and the Alliance for Excellent Education's (2004) *Reading Next* report. External advisors include Margaret Heritage, Ph.D., National Center for Research on Evaluation, Standards, and Student Testing at UCLA; Karin Hess, Ed.D., Center for Assessment (NCIEA); Thomas P. Hogan, Ph.D., University of Scranton; James Milgram, Ph.D., Stanford University; Michael Milone, Ph.D., research psychologist; R. Sharif M. Shakrani, Ph.D., private consultant; Amanda M. VanDerHeyden, Ph.D., private consultant; and James Ysseldyke, Ph.D., University of Minnesota.

Students with a 100-sight-word vocabulary, or who have reached the Probable Reader stage of literacy development in STAR Early Literacy Enterprise, are typically ready to take a STAR Reading Enterprise assessment. STAR Reading Enterprise serves three purposes of particular interest to school and district administrators: (1) to give teachers quick and accurate estimates of students' reading achievement levels, (2) to assess reading achievement relative to national norms, and (3) to provide a means for monitoring growth in a consistent manner longitudinally for all students. Figure 2 shows a sample assessment item.

Teachers who use STAR Reading Enterprise can monitor progress toward college- and career-ready standards, such as the Common Core State Standards, as well as predict proficiency on state tests. After a STAR Enterprise assessment is taken, the software uses the resulting scaled score to locate the student's entry point onto the Core Progress learning progression, helping educators learn more about how the student is performing relative to grade-level expectations. Core Progress provides a road map of skills, spanning from emergent reading to the level of competence required for college and careers, displaying both prerequisite skills students have typically mastered and skills they are ready to develop next.

Figure 2: STAR Reading Enterprise™ Sample Assessment Item

The learning progression, however, is not a straight trajectory. Because students develop at different rates and in different ways, STAR software includes additional resources for targeted instruction, intervention, and enrichment, including Worked Examples, Skill Probes, and Performance Tasks. Additional content will be continuously developed as a means to probe more deeply into students' understandings and skills development (for more about Core Progress, see pp. 10 and 31).

STAR Math Enterprise™ Assessment

STAR Math Enterprise is a challenging, interactive, and brief (about 20 minutes) assessment, consisting of 34 items per test, that evaluates students' mathematical abilities in grades K–12.[3] Like STAR Reading Enterprise, its repeatability and flexibility in administration provide specific advantages for educators:

- Teachers use results from STAR Math Enterprise to facilitate individualized instruction and identify students who most need remediation or enrichment.
- Principals access assessment information through browser-based management and regular, accurate reports on performance at the individual, class, building, and district level.
- Administrators and assessment specialists apply reliable and timely information on mathematical growth at each school and districtwide, which serves as a valid basis for comparing data across schools, grades, and special student populations.

STAR Math Enterprise is a skills-based assessment of math achievement. Table 4 breaks down the STAR Math Enterprise item bank by overall size, number of items administered per testing event, and average administration time.

Table 4: Summary of STAR Math Enterprise™ Item Bank Size and Administration Details

	STAR Math Enterprise
Item Bank Size	More than 5,000
Items Administered per testing event	34 items
Average Administration Time	About 20 minutes

STAR Math Enterprise provides a reliable and valid method for measuring progress towards achievable goals in mathematics. Teachers, principals, literacy coaches, assessment directors, and district-level administrators can use the assessment data for instructional planning, growth measurement, and program evaluation. At an individual student level, STAR can be used for a variety of purposes, including screening, formative assessment, progress monitoring, calculating growth, and outcomes assessment. By using the assessment on a regular basis, such as quarterly or monthly, teachers can monitor progress and make appropriate adjustments to instruction. Research firmly supports progress monitoring, which has shown to be successful in a variety of educational settings.

---
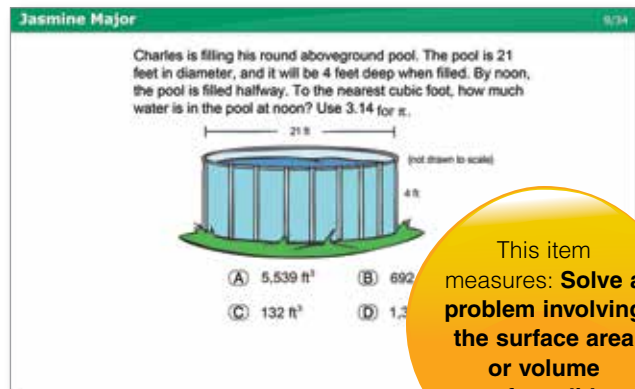
[3] Although STAR Math Enterprise is normed for grades 1–12, kindergarten students may take the assessment with teacher discretion.

As with STAR Reading Enterprise, teachers who use STAR Math Enterprise can monitor progress toward college- and career-ready standards, such as those found in the Common Core State Standards, as well as predict proficiency on state tests. After a STAR Enterprise assessment is taken, the software uses the resulting scaled score to locate the student's entry point onto the Core Progress learning progression, helping educators learn more about how the student is performing relative to grade-level expectations. Core Progress provides a road map of skills, spanning from early numeracy to the level of competence required for college and careers, displaying both prerequisite skills students have typically mastered and skills they are ready to develop next.

The learning progression, however, is not a straight trajectory. Because students develop at different rates and in different ways, the software includes additional resources for targeted instruction, intervention, and enrichment, including Worked Examples, Skill Probes, Performance Tasks, and links to third-party educational resources. Additional content will be continuously developed as a means to probe more deeply into students' understandings and skills development (for more about Core Progress, see pp. 10 and 31).

Students taking a STAR Math Enterprise assessment follow a protocol in which they use blank work paper and pencils during the test administration. As warranted for specific assessment items, the test also provides an onscreen calculator and/or reference sheet. Figure 3 shows a sample assessment item.

Figure 3: STAR Math Enterprise™ Sample Assessment Item

Jasmine Major

Charles is filling his round aboveground pool. The pool is 21 feet in diameter, and it will be 4 feet deep when filled. By noon, the pool is filled halfway. To the nearest cubic foot, how much water is in the pool at noon? Use 3.14 for π.

21 ft

(not drawn to scale)

4 ft

(A) 5,539 ft³   (B) 692
(C) 132 ft³   (D) 1,3

This item measures: **Solve a problem involving the surface area or volume of a solid**

Test Design

Computer adaptive testing (CAT)

STAR Enterprise assessments are computer adaptive tests (CATs). CATs continually adjust the difficulty of each student's test by selecting each assessment item based on the student's previous performance. CATs shorten testing time as well as spare students both the frustration of items that are too difficult and the boredom of items that are too easy.

> "STAR measures offer an important and potentially valuable contribution to RTI."
>
> Shapiro, 2012, p. 20

Decades of research have shown that CATs can be considerably more efficient than conventional tests, which present all students with the same test questions (e.g., Lord, 1980; McBride & Martin, 1983). A well-designed CAT is often two or more times as efficient as a conventional test. For example, to equal the reliability of a 50-item conventional test, a well-designed CAT may use only 25 items to yield the same information in half the time. As noted by Weiss (2004), "Early evidence of improved measurement precision (reliability) and validity (e.g., Johnson & Weiss, 1980; Kingsbury & Weiss, 1980) and large reductions in the number of items administered (typically 50% or more) without having an impact on the psychometric characteristics of test scores for CAT have been confirmed in a number of recent studies (e.g., Mardberg & Carlstedt, 1998; Moreno & Segall, 1997)" (pp. 77–78).

A new line of research suggests that CATs are a sound choice for progress monitoring student performance in response to intervention (RTI) settings. "RTI is a process of providing high quality interventions that are matched to student need, and uses frequent progress monitoring of student response to interventions to assist in making important educational decisions" (Bray & Kehle, 2011, p. 616). Progress monitoring feedback is key to RTI as it tells educators which interventions are helping students most. Thus, "progress-monitoring measures must be frequent, sensitive to instructional change over a short period of time, predictive of overall success as measured by the benchmark assessment, and able to drive instructional decisions" (Shapiro, 2012, p. 9). "STAR measures offer an important and potentially valuable contribution to RTI" (Shapiro, p. 20) in the following ways:

- Frequency of administration—STAR Aassessments were designed to provide educators with flexibility in administering the assessments at the frequency most fitting their needs, whether it be three times per school year for screening, monthly to better understand how student progress is unfolding during the school year with enough time to change the growth trajectory, or as often as weekly for progress monitoring students in tiers 2 and 3 of an RTI framework.

- Sensitivity—STAR Assessments meet all criteria set by the National Center on Intensive Intervention (NCII) (2012a, 2012b, 2012c) and the National Center on Response to Intervention (NCRTI) (2010a, 2010b, 2010c) for "Sensitivity to Student Improvement."

- Predictive power—STAR Assessments meet all criteria set by the NCII and the NCRTI for "Predictive Validity of the Slope of Improvement," as well as criteria set by the NCRTI for "Classification Accuracy." In addition, a (2012) study found that STAR Math "was the single best predictor of PSSA scores across grades" (Shapiro & Gebhardt, p. 303) when compared to CBM measures. For additional predictive validity evidence for each STAR, see Psychometric Properties, p. 19.

- Impact on instructional decisions—STAR Assessments meet all criteria set by NCII and NCRTI for both "Decision Rules for Changing Instruction" and "Decision Rules for Increasing Goals." Core Progress learning progressions—which place students scores within a progression of learning—make the data from STAR Assessments immediately actionable and facilitate instructional planning (for more information, see Instructional planning with Core Progress, p. 31).

Item response theory and its role in CAT

Tailoring item difficulty to match a student's knowledge or skill level can be done in a number of different ways; however, most CATs use item response theory (IRT) as the basis for both adaptive item selection and test scoring. IRT puts student performance and item difficulty on the same scale and offers a means to estimate the probability that a student will answer a given test item correctly. IRT models provide a way to measure each item's degree of difficulty and to estimate each student's achievement level from the pattern of correct and incorrect responses to items.

With item response theory, scientists can calculate the probability of a correct response to an item as a function of student ability. As student ability increases, so does the probability the student will answer correctly. Additionally, because some test items are harder than others, the probability trend differs from one item to another. Figure 4 shows the probability functions for three test items: one that's easy, one that's moderately difficult, and one that's very difficult.

Figure 4: Illustration of a Student's Reactions to Three Test Items of Varying Difficulty
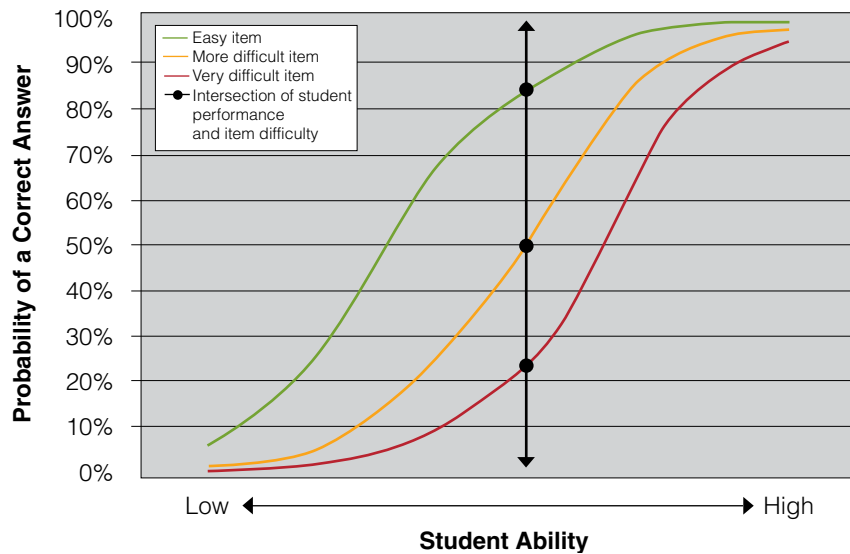
During a STAR Enterprise assessment administration, the software automatically moves up or down the item scale to select questions based on a student's answers. If the student answers a question correctly, the next question will be more difficult. If the student answers incorrectly, the next question will be less difficult. Unlike manual paper-and-pencil assessments, STAR Enterprise assessments dynamically adjust to each student's unique responses. As a result, STAR pinpoints student achievement levels quickly and efficiently.

Figure 5 displays an example progression of less difficult and more challenging items based on a student's previous item responses during a CAT administration. It also shows how selecting items tailored to a student's ability helps to reduce measurement error as the test progresses.

Figure 5: How Computer-Adaptive Technology Works
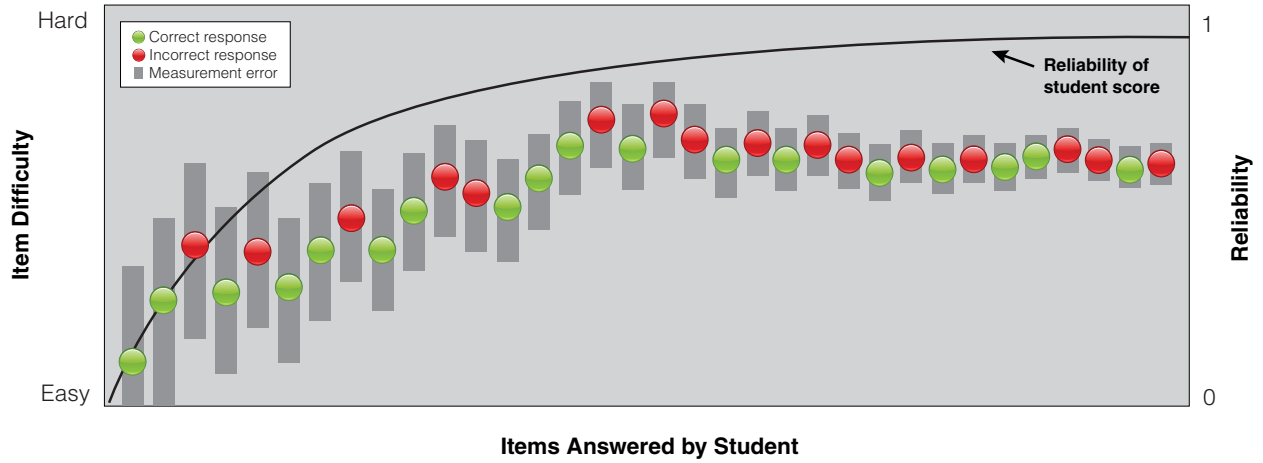
Core Progress™ Learning Progressions—The Bridge Between Assessment and Instruction

A learning progression is a continuum of expected learning, beginning with emergent reading or early numeracy skills and progressing to the level of competence required for college and careers. The skills are interconnected and related, formed from requisites and prerequisites, and represent how students typically advance their learning in a subject area. According to Heritage (2008), "Learning progressions that clearly articulate a progression of learning in a domain can provide the big picture of what is to be learned, support instructional planning, and act as a touchstone for formative assessment" (p. 1).

> In July 2013, Renaissance Learning released two new learning progressions built specifically for the Common Core.

Skills in a learning progression are not meant to be taught sequentially; rather, a student's placement on a learning progression begins with a student's score from a standardized test of achievement. This information helps orient student and teacher to where the student has been, where the student is headed, and the skills with which they may need guidance in order to arrive at their destination successfully.

Evolution of Core Progress™

To build a bridge between assessment and instruction, Renaissance Learning created the Core Progress for Reading and Core Progress for Math learning progressions. Members of the Renaissance Learning standards team rigorously developed, tested, and validated Core Progress. For both reading and math, standards experts identified the initial order of item difficulty by researching reading and math theory, examining widely accepted frameworks such as state standards, reviewing the Common Core State Standards (CCSS), and consulting nationally recognized reading and math experts.

The road map of skills in Core Progress helps teachers monitor progress toward college- and career-ready standards. Using a student's STAR scaled score, Core Progress displays student progress in skills relative to grade-level expectations.

All students follow individual paths to achieve personalized goals. Because students develop reading and math ability at different rates and in different ways, a student's progression through Core Progress does not follow a straight trajectory. Additional resources, such as Worked Examples, Skill Probes, Performance Tasks, and links to third-party educational resources, help teachers meet students at their individual achievement levels for targeted instruction, intervention, and enrichment.

Built for the Common Core State Standards

As the majority of states implemented the Common Core State Standards, Renaissance Learning recognized a need for learning progressions created expressly for these new standards. In July 2013, Renaissance Learning released two new learning progressions built specifically for the CCSS:

- **Core Progress™ Learning Progression for Reading—Built for the Common Core State Standards**
- **Core Progress™ Learning Progression for Math—Built for the Common Core State Standards**

Like the original Core Progress, the new CCSS-specific learning progressions present a continuum of skills from emergent reading and early numeracy through the level of knowledge required for college and careers, as well as display both prerequisite skills students have mastered and skills they are ready to develop next. The new learning progressions are different in that they were built, from the ground up, specifically for the Common Core State Standards.

Renaissance Learning standards experts began this process with a close analysis of the CCSS, identifying each standard's inherent skills, intent, and key terminology. They also immersed themselves in the literature and resources available regarding the CCSS to determine how the standards were being interpreted and implemented by states and relevant consortia. All of this ensured that the new learning progressions included incremental steps of learning to fulfill the intent of the standards and ultimately culminate in college and career readiness.

Path from test blueprint to learning progression

Empirical testing has found a strong statistical link between the progression of skills in Core Progress and the assessed difficulty level of STAR Enterprise test items, meaning educators can use scores from the assessments to identify both what a student knows and what they need to work on. As Figure 6 shows, a STAR assessment's blueprint working in tandem with CAT technology ultimately dictates which items are presented to each student. While each STAR test event is unique, the blueprint ensures that a certain number of items from the domains and skill sets are presented to each student.

Figure 6: How it Works: From STAR™ Test Blueprint to Core Progress™ Learning Progression
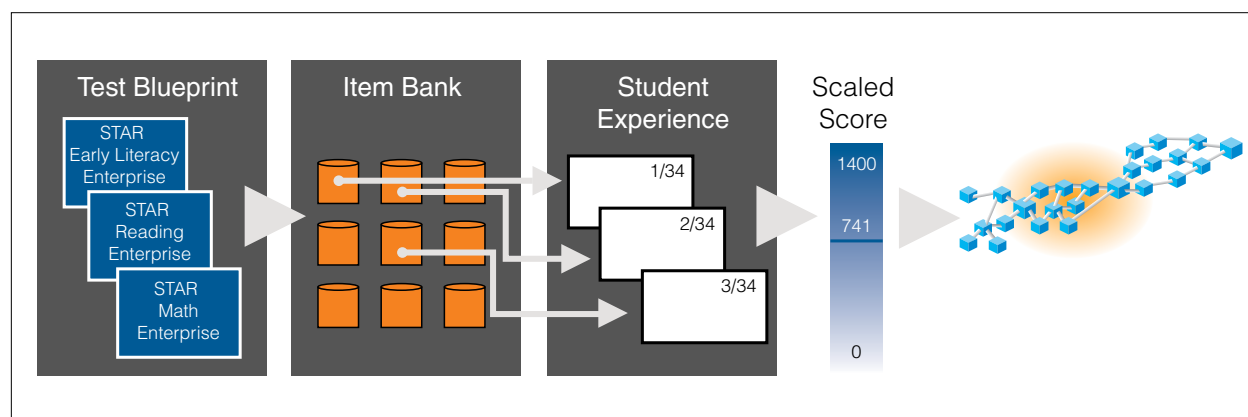
Depending on the state in which you reside, you will either have access to the original Core Progress learning progression or the Core Progress Learning Progression—Built for the Common Core State Standards.

After a student takes a STAR Enterprise assessment, the software uses the resulting scaled score to find the student's entry point onto the Core Progress learning progression and then reports the skills the student has likely mastered in prior grades and those the student is ready to develop next, helping teachers to focus instruction. For more information about how Core Progress helps tailor student instruction, see Instructional planning with Core Progress, p. 31.[4]

---

[4] For more in-depth information, please see:
  Core Progress for Reading: Empirically Validated Learning Progressions (http://doc.renlearn.com/KMNet/R0053985FA6D567F.pdf)
  Core Progress for Math: Empirically Validated Learning Progressions (http://doc.renlearn.com/KMNet/R00552482161352C.pdf)

Skills in Core Progress™ Learning Progression–Built for the Common Core State Standards
The order of skills presented in the new learning progressions built for the CCSS emerged from Renaissance Learning content experts' deep study of the standards.

Figure 7 displays the organization of the domains and skill areas in the learning progression for early literacy.

Figure 7: Core Progress™ Learning Progression for Reading—Built for the Common Core State Standards: Domains and Skill Areas (Early Literacy)

Foundational Skills

Print Concepts
- Directionality
- Letters and Words
- Word Length
- Word Borders
- Visual Discrimination / Alphabetic Principle
- Alphabetic Sequence
- Print Features

Phonological Awareness
- Rhyming and Word Families
- Blending, Counting, and Segmenting Syllables
- Blending and Segmenting
- Distinguishing between Long and Short Vowel Sounds
- Isolating Initial, Final, and Medial Phonemes
- Adding/Substituting Phonemes

Phonics and Word Recognition
- Spelling-Sound Correspondences: Consonants
- Spelling-Sound Correspondences: Vowels
- Regular and Irregular Spellings / High-Frequency words
- Inflectional Endings / Affixes
- Syllables

Fluency
- Purpose of Reading / Reading with Comprehension
- Reading Rate WCPM
- Prosody
- Repair Strategies

Language

Vocabulary Acquisition and Use
- Real-Life Word Connections and Applications
- Word Reference Materials
- Antonyms
- Synonyms
- Structural Analysis
- Word Relationships
- Context Clues
- Vocabulary in Context
- Multiple-Meaning Words
- Figures of Speech
- Connotation

As Figure 8 shows, for reading, the organization of the learning progression reflects the CCSS with four domains: (1) Foundational Skills, (2) Language, (3) Literature, and (4) Informational Text (which reflects the emphasis on nonfiction text in the standards).

Figure 8: Core Progress™ Learning Progression for Reading—Built for the Common Core State Standards: Domains and Skill Areas

Foundational Skills

Print Concepts
- Directionality
- Letters and Words
- Word Length
- Word Borders
- Visual Discrimination / Alphabetic Principle
- Alphabetic Sequence
- Print Features

Phonological Awareness
- Rhyming and Word Families
- Blending, Counting, and Segmenting Syllables
- Blending and Segmenting
- Distinguishing between Long and Short Vowel Sounds
- Isolating Initial, Final, and Medial Phonemes
- Adding/Substituting Phonemes

Phonics and Word Recognition
- Spelling-Sound Correspondences: Consonants
- Spelling-Sound Correspondences: Vowels
- Regular and Irregular Spellings / High-Frequency words
- Inflectional Endings / Affixes
- Syllables

Fluency
- Purpose of Reading / Reading with Comprehension
- Reading Rate WCPM
- Prosody
- Repair Strategies

Language

Vocabulary Acquisition and Use
- Real-Life Word Connections and Applications
- Word Reference Materials
- Antonyms
- Synonyms
- Structural Analysis
- Word Relationships
- Context Clues
- Vocabulary in Context
- Multiple-Meaning Words
- Figures of Speech
- Connotation

Literature

Key Ideas and Details
- Character
- Setting
- Plot
- Theme
- Summary
- Inference and Evidence

Craft and Structure
- Point of View
- Structure of Literary Text
- Word Meaning
- Author's Word Choice and Figurative Language
- Connotation

Integration of Knowledge and Ideas
- Modes of Representation
- Analysis and Comparison

Range of Reading and Level of Text Complexity
- Range of Reading
- Development of Independence

Informational Text

Key Ideas and Details
- Main Idea and Details
- Inference and Evidence
- Prediction
- Sequence
- Compare and Contrast
- Cause and Effect
- Summary
- Connections and Relationships

Craft and Structure
- Text Features
- Author's Purpose and Perspective
- Word Meaning
- Connotation
- Organization
- Author's Word Choice and Figurative Language

Integration of Knowledge and Ideas
- Modes of Representation
- Argumentation
- Analysis and Comparison

Range of Reading and Level of Text Complexity
- Range of Reading
- Development of Independence

In Figures 9 and 10, the organization of the learning progression for math is identical to the CCSS framework for grades K–8 and high school.

Figure 9: Core Progress™ Learning Progression for Math—Built for the Common Core State Standards: Domains and Skill Areas (K–8)

Counting and Cardinality

- Whole Numbers: Counting, Comparing, and Ordering

Operations and Algebraic Thinking

- Algebraic Thinking
- Evaluate Numerical Expressions
- Whole Numbers: Addition and Subtraction
- Whole Numbers: Counting, Comparing, and Ordering
- Whole Numbers: Multiplication and Division

Number and Operations in Base Ten

- Decimal Concepts and Operations
- Powers, Roots, and Radicals
- Whole Numbers: Addition and Subtraction
- Whole Numbers: Counting, Comparing, and Ordering
- Whole Numbers: Multiplication and Division
- Whole Numbers: Place Value

Number and Operations — Fractions

- Decimal Concepts and Operations
- Fraction Concepts and Operations

Ratios and Proportional Relationships

- Percents, Ratios, and Proportions

The Number System

- Coordinate Geometry
- Decimal Concepts and Operations
- Fraction Concepts and Operations
- Integers
- Whole Numbers: Multiplication and Division

Expressions and Equations

- Evaluate and Use Variable Expressions
- Evaluate Numerical Expressions
- Linear Equations and Inequalities
- Powers, Roots, and Radicals
- Quadratic and Nonlinear Equations and Inequalities
- Systems of Equations and Inequalities

Functions

- Relations and Functions

Geometry

- Angles, Segments, and Lines
- Congruence and Similarity
- Coordinate Geometry

- Fraction Concepts and Operations

- Geometry: Three-Dimensional Shapes and Attributes
- Geometry: Two-Dimensional Shapes and Attributes
- Perimeter, Circumference, and Area
- Right Triangles and Trigonometry
- Surface Area and Volume
- Transformations

Measurement and Data

- Angles, Segments, and Lines
- Data Representation and Analysis
- Geometry: Two-Dimensional Shapes and Attributes
- Measurement
- Money
- Perimeter, Circumference, and Area
- Surface Area and Volume
- Time
- Whole Numbers: Addition and Subtraction
- Whole Numbers: Counting, Comparing, and Ordering

Statistics and Probability

- Combinatorics and Probability
- Data Representation and Analysis

Figure 10: Core Progress™ Learning Progression for Math—Built for the Common Core State Standards: Domains and Skill Areas (High School)

The Real Number System

- Fraction Concepts and Operations
- Powers, Roots, and Radicals

Quantities

- Data Representation and Analysis

Seeing Structure in Expressions

- Algebra of Polynomials
- Linear Equations and Inequalities
- Quadratic and Nonlinear Equations and Inequalities
- Relations and Functions

Arithmetic with Polynomials and Rational Expressions

- Algebra of Polynomials

Creating Equations

- Linear Equations and Inequalities

Reasoning with Equations and Inequalities

- Linear Equations and Inequalities
- Quadratic and Nonlinear Equations and Inequalities
- Relations and Functions
- Systems of Equations and Inequalities

Interpreting Functions

- Relations and Functions

Building Functions

- Relations and Functions

Linear, Quadratic, and Exponential Models

- Linear Equations and Inequalities
- Quadratic and Nonlinear Equations and Inequalities

The Complex Number System

- Algebra of Polynomials
- Complex Numbers

Trigonometric Functions

- Right Triangles and Trigonometry

Congruence

- Angles, Segments, and Lines
- Congruence and Similarity
- Geometry: Two-Dimensional Shapes and Attributes
- Polygons and Circles
- Transformations

Similarity, Right Triangles, and Trigonometry

- Congruence and Similarity
- Right Triangles and Trigonometry
- Transformations

Circles

- Polygons and Circles

Expressing Geometric Properties with Equations

- Coordinate Geometry
- Polygons and Circles

Geometric Measure and Dimension

- Geometry: Three-Dimensional Shapes and Attributes
- Perimeter, Circumference, and Area
- Surface Area and Volume

Modeling with Geometry

- Coordinate Geometry
- Geometry: Three-Dimensional Shapes and Attributes
- Perimeter, Circumference, and Area
- Polygons and Circles
- Right Triangles and Trigonometry
- Surface Area and Volume

Conditional Probability and the Rules of Probability

- Combinatorics and Probability

Using Probability to Make Decisions

- Combinatorics and Probability

Interpreting Categorical and Quantitative Data

- Data Representation and Analysis

Making Inferences and Justifying Conclusions

- Data Representation and Analysis

Skills in original Core Progress™ Learning Progression
Development of the original Core Progress learning progressions for reading and math took into account research as well as state and other standards.

Figure 11 shows the organization of the early literacy and early numeracy skills in the learning progression within three key domains: (1) Word Knowledge and Skills, (2) Comprehension Strategies and Constructing Meaning, and (3) Numbers and Operations.

Figure 11: Core Progress™ for Reading Learning Progression: Domains and Skill Sets (Early Literacy)

Word Knowledge and Skills

Alphabetic Principle
- Alphabetic Knowledge
- Alphabetic Sequence
- Letter Sounds

Concept of Word
- Print Concepts: Word Length
- Print Concepts: Word Borders
- Print Concepts: Letters and Words

Visual Discrimination
- Letters
- Identification and Word Matching

Phonemic Awareness
- Rhyming and Word Families
- Blending Word Parts
- Blending Phonemes
- Initial and Final Phonemes
- Consonant Blends (PA)
- Medial Phoneme Discrimination
- Phoneme Segmentation
- Phoneme Isolation/Manipulation

Phonics
- Short Vowel Sounds
- Initial Consonant Sounds
- Final Consonant Sounds
- Long Vowel Sounds
- Variant Vowel Sounds
- Consonant Blends (PH)
- Consonant Digraphs
- Other Vowel Sounds
- Sound-Symbol Correspondence: Consonants
- Word Building
- Sound-Symbol Correspondence: Vowels
- Word Families/Rhyming

Structural Analysis
- Words with Affixes
- Syllabification
- Compound Words

Vocabulary
- Word Facility
- Synonyms
- Antonyms

Comprehension Strategies and Constructing Meaning

Sentence-level Comprehension
- Comprehension at the Sentence Level

Paragraph-level Comprehension
- Comprehension of Paragraphs

Numbers and Operations

Early Numeracy
- Number Naming and Number Identification
- Number Object Correspondence
- Sequence Completion
- Composing and Decomposing
- Measurement

In Figure 12, for reading, the learning progression is organized by five domains: (1) Word Knowledge and Skills, (2) Comprehension Strategies and Constructing Meaning, (3) Understanding Author's Craft, (4) Analyzing Literary Text, and (5) Analyzing Argument and Evaluating Text.

Figure 12: Core Progress™ for Reading Learning Progression: Domains and Skills

Word Knowledge and Skills

Vocabulary Strategies
- Use context clues
- Use structural analysis

Vocabulary Knowledge
- Recognize and understand synonyms
- Recognize and understand homonyms and multi-meaning words
- Recognize connotation and denotation
- Understand idioms
- Understand analogies

Analyzing Literary Text

Literary Elements
- Identify and understand elements of plot
- Identify and understand setting
- Identify characters and understand characterization
- Identify and understand theme
- Identify the narrator and point of view

Genre Characteristics
- Identify fiction and nonfiction, reality and fantasy
- Identify and understand characteristics of genres

Understanding Author's Craft

Author's Choices
- Understand figurative language
- Understand literary devices
- Identify sensory detail

Comprehension Strategies and Constructing Meaning

Reading Process Skills
- Make predictions
- Identify author's purpose
- Identify and understand text features
- Recognize an accurate summary of text
- Use repair strategies

Constructing Meaning
- Understand vocabulary in context
- Draw conclusions

- Identify and understand main ideas
- Identify details
- Extend meaning or form generalizations
- Identify and differentiate fact and opinion

Organizational Structure
- Identify organizational structure
- Understand cause and effect
- Understand comparison and contrast
- Identify and understand sequence

Analyzing Argument and Evaluating Text

Analysis
- Identify bias and analyze text for logical fallacies
- Identify and understand persuasion

Evaluation
- Evaluate reasoning and support
- Evaluate credibility

Figure 13 shows the math learning progression's organization within four domains: (1) Numbers and Operations, (2) Algebra (3) Geometry and Measurement, (4) Data Analysis, Statistics, and Probability.

Figure 13: Core Progress™ for Math Learning Progression: Domains and Skill Sets

Numbers and Operations

- Count with objects and numbers
- Identify odd and even numbers
- Relate place and value to a whole number
- Add and subtract whole numbers without regrouping
- Add and subtract whole numbers with regrouping
- Multiply whole numbers
- Divide whole numbers without a remainder in the quotient
- Divide whole numbers with a remainder in the quotient
- Identify, compare, and order fractions
- Add and subtract fractions with like denominators
- Find prime factors, common factors, and common multiples
- Add and subtract fractions with unlike denominators
- Convert between an improper fraction and a mixed number
- Relate a decimal to a fraction
- Relate place and value to a decimal number
- Add or subtract decimal numbers
- Divide a whole number resulting in a decimal quotient
- Multiply and divide with fractions

- Multiply and divide with decimals
- Relate a decimal number to a percent
- Solve a proportion, rate, or ratio
- Evaluate a numerical expression
- Perform operations with integers
- Determine a square root
- Solve a problem involving percents

Data Analysis, Statistics, and Probability

- Read or answer a question about charts, tables, or graphs
- Use a chart, table, or graph to represent data
- Determine a measure of central tendency
- Use a proportion to make an estimate
- Determine the probability of one or more events

Algebra

- Relate a rule to a pattern
- Determine the operation given a situation
- Graph on a coordinate plane
- Evaluate an algebraic expression or function
- Solve a linear equation

- Determine a linear equation
- Identify characteristics of a linear equation or function
- Solve a system of linear equations
- Determine a system of linear equations
- Simplify an algebraic expression
- Solve a linear inequality
- Solve a nonlinear equation
- Graph a 1-variable inequality

Geometry and Measurement

- Relate money to symbols, words, and amounts
- Use the vocabulary of geometry and measurement
- Determine a missing figure in a pattern
- Determine a measurement
- Tell time
- Calculate elapsed time
- Solve a problem involving the perimeter of a shape
- Solve a problem involving the area of a shape
- Identify congruence and similarity of geometric shapes
- Solve a problem involving the surface area or volume of a solid
- Determine a missing measure or dimension of a shape

Psychometric Properties

The computer-adaptive STAR Assessments are highly rated for reliability and validity by key federal groups, such as the National Center on Intensive Intervention, the National Center on Response to Intervention, and the National Center on Student Progress Monitoring.

In 2012, STAR Assessments were highly rated for progress monitoring by the federally funded National Center on Intensive Intervention (NCII), whose mission is "to build state and district capacity to support educators in using data-based individualization to effectively implement intensive interventions in reading, mathematics, and behavior in Grades K–12" (http://www.intensiveintervention.org), in the organization's first review of progress-monitoring tools.

Earlier, in 2009, the U.S. Department of Education began funding the National Center on Response to Intervention (NCRTI), whose mission is "to provide technical assistance to states and districts and building the capacity of states to assist districts in implementing proven models for RTI/EIS" (www.rti4success.org). That same year, STAR Early Literacy, STAR Reading, and STAR Math were among the first assessments highly rated by the NCRTI for screening and progress monitoring. In subsequent reviews, STAR Assessments have maintained strong ratings, meaning they fulfill both these key elements of a school's RTI framework. For information on using STAR Enterprise assessments in intervention settings, see Pupose and Frequency, p. 28.

> STAR Assessments are highly rated for reliability and validity by key federal groups, such as the National Center on Intensive Intervention, the National Center on Response to Intervention, and the National Center on Student Progress Monitoring.

STAR Assessments have received high marks as tools for Response to Intervention since 2006 when the NCRTI's predecessor, the National Center on Student Progress Monitoring, first deemed STAR Early Literacy, STAR Reading, and STAR Math reliable and valid for progress monitoring (http://www.studentprogress.org/chart/docs/print_chart122007.pdf).

Each STAR assessment followed a unique path to determine reliability and validity, which is explained below along with lists of the wide range of assessments to which each STAR assessment relates.

Reliability and validity of STAR Early Literacy Enterprise™

Reliability
Test reliability is often described as a measure of the consistency of test scores; tests must yield somewhat consistent results in order to be useful. Two kinds of consistency are of concern when evaluating a test's measurement precision: internal consistency and the consistency of the scores obtained when an assessment is given two or more times.

The internal consistency of STAR Early Literacy Enterprise assessments has been calculated using a method referred to as *generic reliability*, which uses the conditional measurement error of individual students' tests to estimate what percentage of the variation in STAR test scores is attributable to the attribute the test is intended to measure. Consistency of scores across multiple administrations of the assessment to the same students is measured by *retest reliability*, which is the coefficient of correlation between pairs of test scores earned by the same students on different occasions.

The generic estimates of internal consistency reliability were calculated from analyes of the test scores and their estimated conditional measurement error in a balanced random sample of 10,000 students in each grade, pre-K through 3, who took STAR Early Literacy Enterprise in fall 2012. Another random sample of students who took SEL Enterprise two or more times within a 2-week period across the same school year was analyzed in to order to calculate retest reliability. Table 5 displays both the internal consistency and the retest reliability estimates, by grade and for the five grades combined. The combined-grades reliability coefficients are 0.85 for internal consistency, and 0.79 for consistency on retest.

Table 5: Internal Consistency and Retest Reliability of STAR Early Literacy Enterprise™
Assessments Taken Between June 2012 and June 2013

Grade	Internal Consistency		Retest Reliability	
	Students	Reliability Coefficient	Students	Reliability Coefficient
All	3,083,334	0.85	25,000	0.79
Pre-K	54,144	0.81	5,000	0.59
K	1,427,660	0.80	5,000	0.50
1	1,187,216	0.82	5,000	0.47
2	340,912	0.85	5,000	0.64
3	73,402	0.89	5,000	0.74

Validity

Evidence of the validity of any educational assessment has a number of facets that, in aggregate, constitute empirical support for the use of the assessments for specific purposes, and for the inferences that are to be made on the basis of students' test scores. A crucial facet is the content of the tests; content-related evidence of validity lies in the degree of correspondence, or alignment, between the knowledge and skills measured by an assessment's test items and the knowledge and skills intended to be taught and learned in a given curriculum at a given grade level or levels. STAR Early Literacy Enterprise content is aligned to curriculum standards at the state and national levels—including the Common Core State Standards (see Core Progress Learning Progressions—The Bridge Between Assessment and Instruction, p. 10).

It could be argued that solid evidence of psychometric reliability, combined with a high degree of alignment of test content to curriculum standards, is evidence enough of an assessment's validity. However, a number of other measures complement or corroborate those two facets and serve to further strengthen an assessment's claims of validity: cumulative evidence of criterion-related validity, convergent and discriminant validity evidence, demonstrated accuracy of screening and diagnostic classifications, among others. Altogether, these are among the components of *construct validity*, in other words, evidence the assessments measure specific attributes as claimed and are appropriate for specific uses and inferences. Construct validity evidence is cumulative in nature; when first released, an assessment may have sound evidence that is consistent with construct validity, but over time additional evidence may and should be accumulated and documented.

To support, STAR Early Literacy as a measure of literacy skills, Renaissance Learning knew it was necessary that its scores correlate highly with other measures of reading, literacy, and readiness. To evaluate this, Renaissance Learning performed a multifaceted validity research study of STAR Early Literacy prior to the assessment's initial release to assess reliability, criterion-related validity, and score distributions by age and grade. The participating school districts, specific schools, and individual students were

approximately representative of the U.S. school population in terms of geographic region, school system and per-grade district enrollment, and socioeconomic status. The final study sample included approximately 11,000 students from 84 schools in the U.S. and Canada.

Renaissance Learning asked teachers participating in the study to submit student scores from other assessments of reading, early literacy, readiness, and social skills. Scores were received for more than 2,400 students. The resulting correlation estimates were substantial and reflect well on the concurrent validity of STAR Early Literacy as a tool for assessing early literacy skills. Subsequent to the original validity study, a number of additional studies, including both concurrent and predictive correlational studies, studies of classification accuracy, and others, have been conducted. Table 6 summarizes the results of more than 80 concurrent and predictive validity studies conducted for STAR Early Literacy. The average correlations observed in these studies range from 0.52 to 0.77; correlations in that range are considered moderate to strong. Below the table is a list of major assessments of early literacy skills that have been found to correlate well with scores on STAR Early Literacy.

Table 6: Summary of STAR Early Literacy™ Validity Studies

Grade	Predictive			Concurrent		
	Studies	Students	Average Correlation	Studies	Students	Average Correlation
K	15	30,423	0.52	6	198	0.64
1	15	24,525	0.62	7	281	0.68
2	15	5,370	0.67	12	513	0.52
3	2	558	0.67	9	384	0.57

STAR Early Literacy™ relates to several assessments of early literacy skills
Studies have been conducted with STAR Early Literacy and the following assessments to correlate the tests:

• AIMSweb
• Alabama Early Learning Inventory
• Brigance K & 1 Screen for Kindergarten and First Grade Children
• Canadian Achievement Test
• Child Observation Record (COR)
• Developing Skills Checklist (DSC)
• Developmental Indicators for the Assessment of Learning (DIAL-3)
• Dynamic Indicators of Basic Early Literacy Skills (DIBELS)
• easyCBM
• Florida Comprehensive Assessment Test (FCAT)
• Gates-MacGinitie Reading Test (GMRT)
• Group Reading Assessment and Diagnostic Evaluation (GRADE)
• Indiana Statewide Testing for Educational Progress (ISTEP)
• Iowa Test of Basic Skills (ITBS)
• Kaufman Survey of Early Academic and Language Skills (K-SEALS)
• Metropolitan Early Childhood Assessment Program (MKIDS)
• Metropolitan Readiness Test (MRT)
• Michigan Literacy Progress Profile (MLPP)
• NWEA Levels Test
• Running Records

- Stanford Achievement Test (SAT-9)
- Stanford Test of Academic Skills
- TerraNova
- Test of Phonological Awareness (TOPA)
- Texas Primary Reading Inventory (TPRI)
- Woodcock Reading Mastery Tests-Revised/Normative Update

Reliability and validity of STAR Reading Enterprise™

Reliability

The reliability of STAR Reading Enterprise assessments was estimated using two methods, internal consistency (generic reliability coefficients) and test-retest correlation coefficients, in a random national sample of more than 1.2 million STAR Reading Enterprise tests administered between September 2012 and June 2013. The retest correlation coefficients were based on samples of 5,000 students per grade, from the same dataset. Results are displayed in Table 7. The internal consistency reliability estimates were very high, equaling or exceeding those of most major published assessments. Over all grades combined, the reliability was 0.97; it ranged from 0.93 to 0.95 within grades. Retest reliability estimates were 0.90 for all grades combined, and ranged from 0.54 to 0.85 within grades.

Table 7: Internal Consistency and Retest Reliability of STAR Reading Enterprise™
Assessments Taken Between June 2012 and June 2013

	Internal Consistency		Retest Reliability	
	Students	**Reliability Coefficient**	**Students**	**Reliability Coefficient**
All	1,227,915	0.97	60,000	0.90
1	100,000	0.95	5,000	0.54
2	100,000	0.94	5,000	0.66
3	100,000	0.94	5,000	0.75
4	100,000	0.93	5,000	0.77
5	100,000	0.93	5,000	0.78
6	100,000	0.93	5,000	0.83
7	100,000	0.94	5,000	0.82
8	100,000	0.94	5,000	0.83
9	95,171	0.94	5,000	0.85
10	94,624	0.95	5,000	0.85
11	93,118	0.95	5,000	0.85
12	89,031	0.95	5,000	0.85

Validity

As noted in the discussion of STAR Early Literacy validity, content is a crucial facet of test validity; content-related evidence of validity lies in the degree of correspondence, or alignment, between the knowledge and skills measured by an assessment's test items and the knowledge and skills intended to be taught and learned in a given curriculum at a given grade level or levels. STAR Reading Enterprise content is aligned to curriculum standards at the state and national levels—including the Common Core State Standards (see Core Progress Learning Progressions—The Bridge Between Assessment and Instruction, p. 10).

Psychometric reliability, combined with a high degree of alignment of test content to curriculum standards may be evidence enough of an assessment's validity. However, other measures complement or corroborate those two facets and serve to further strengthen an assessment's claims of validity.

To support STAR Reading Enterprise as a measure of both reading comprehension and a broad range of other reading skills, Renaissance Learning has collected a wide range of correlations between scores on STAR Reading and scores on other recognized, established measures of different aspects of reading achievement, such as survey achievement tests, diagnostic reading measures, and state accountability tests, among others. Table 8 summarizes the results of more than 400 concurrent and predictive validity studies conducted for STAR Reading, involving a total of more than 1 million students. The average correlations observed in these studies range from 0.60 to 0.87; correlations in that range are considered strong. Below the table is a list of state assessments that have been found to correlate well with scores on STAR Reading.

Table 8: Summary of STAR Reading™ Validity Studies

Grade	Predictive			Concurrent and Other External Validity		
	Studies	Students	Average Correlation	Studies	Students	Average Correlation
1	6	74,77	.68	15	1,135	.77
2	10	184,434	.78	32	4,142	.72
3	30	200,929	.80	44	4,051	.75
4	25	185,528	.82	41	5,409	.75
5	29	126,029	.82	40	3,588	.75
6	23	82,189	.82	37	2,728	.71
7	23	64,978	.81	33	3,294	.70
8	25	34,764	.81	29	2,148	.72
9	8	9,567	.83	15	949	.72
10	9	7,021	.85	11	566	.61
11	6	6,653	.86	6	324	.70
12	2	3,107	.86	4	165	.74

STAR Reading™ relates to several state assessments
Studies have been conducted with STAR Reading and the following assessments to statistically link[5] the tests:

• ACT EXPLORE
• Alabama Reading and Mathematics Test+ (ARMT+)
• Alaska's Standards Based Assessment (SBA)
• Arizona's Instrument to Measure Standards (AIMS)
• Arkansas Augmented Benchmark Examinations (AABE)
• California Standards Tests (CST)
• Colorado—Transitional Colorado Assessment Program (TCAP)
• Connecticut Mastery Test (CMT4)
• Delaware Comprehensive Assessment System (DCAS)
• Florida Comprehensive Assessment Test 2.0 (FCAT 2.0)

---

[5] Statistical linking studies are continuously underway and Performance Reports for new states are released on a regular basis. For the most up-to-date list of state assessments with statistical links to STAR Assessments, email research@renlearn.com. Technical manuals are also available upon request and include information on the assessments that correlate with STAR Assessments.

- Georgia's Criterion-Referenced Competency Tests (CRCT)
- Idaho Standards Achievement Tests (ISAT)
- Illinois Standards Achievement Test (ISAT)
- Indiana Reading Evaluation and Determination (IREAD-3)
- Indiana Statewide Testing for Education Progress-Plus (ISTEP+) Assessments
- Iowa Assessment (IA)
- Kansas State Assessment Program (KSAP)
- Kentucky Performance Rating for Educational Progress (K-PREP) Tests
- Louisiana Educational Assessment Program (LEAP) and Integrated Educational Assessment Program (iLEAP) Assessments
- Maine—New England Common Assessment Program (NECAP)
- Massachusetts Comprehensive Assessment System (MCAS)
- Michigan Educational Assessment Program (MEAP)
- Minnesota Comprehensive Assessments (MCAs)
- Mississippi Curriculum Test, Second Edition (MCT2)
- Missouri Assessment Program (MAP) Grade-Level Assessments
- Montana's Criterion-Referenced Test (CRT)
- Nebraska State Accountability (NeSA) Reading Test
- Nevada's Criterion-Referenced Test (CRT)
- New Hampshire—New England Common Assessment Program (NECAP)
- New Jersey Assessment of Skills and Knowledge (NJ ASK)
- New Mexico Standards Based Assessments (SBA)
- New York State Assessment Program (NYSTP)
- North Carolina End-of-Grade (NC EOG) Tests
- North Dakota State Assessment (NDSA)
- Ohio Achievement Assessments (OAA)
- Oklahoma Core Curriculum Tests (OCCT)
- Pennsylvania's System of School Assessment (PSSA)
- Rhode Island—New England Common Assessment Program (NECAP)
- South Carolina Palmetto Assessment of State Standards (SCPASS)
- South Dakota State Test of Educational Progress (DSTEP)
- Tennessee Comprehensive Assessment Program (TCAP)
- Texas—State of Texas Assessments of Academic Readiness (STAAR)
- Utah's Criterion-Referenced Test for English Language Arts
- Vermont—New England Common Assessment Program (NECAP)
- Virginia Standards of Learning (SOL)
- Washington—Measurements of Student Progress (MSP)
- West Virginia Educational Standards Test 2 (WESTEST 2)
- Wisconsin Knowledge and Concepts Examination (WKCE)
- Wyoming—Proficiency Assessments for Wyoming Students (PAWS)

Reliability and validity of STAR Math Enterprise™

Reliability

The reliability of STAR Math Enterprise assessments was estimated using two methods, internal consistency (generic reliability coefficients) and test-retest correlation coefficients, in a national sample of more than 9 million STAR Math Enterprise tests administered between September 2012 and June 2013. The retest correlation coefficients were based on random samples of 5,000 students per grade from the same dataset. Results are displayed in Table 9. The internal consistency reliability estimates were very high, equaling or exceeding those of most major published assessments. Over all grades combined, the reliability was 0.97; it ranged from 0.90 to 0.95 within grades. Retest reliability estimates were 0.93 for all grades combined, and ranged from 0.76 to 0.84 within grades.

Table 9: Internal Consistency and Retest Reliability of STAR Math Enterprise™
Assessments Taken Between June 2012 and June 2013

Validity

Grade	Internal Consistency		Retest Reliability	
	Students	Reliability Coefficient	Students	Reliability Coefficient
All	9,311,595	0.97	60,000	0.93
1	805,980	0.90	5,000	0.76
2	1,254,611	0.91	5,000	0.80
3	1,330,600	0.92	5,000	0.81
4	1,306,386	0.92	5,000	0.83
5	1,227,139	0.93	5,000	0.83
6	968,367	0.93	5,000	0.84
7	785,789	0.94	5,000	0.82
8	721,994	0.94	5,000	0.83
9	327,455	0.93	5,000	0.83
10	241,728	0.94	5,000	0.82
11	167,902	0.94	5,000	0.83
12	108,492	0.95	5,000	0.80

As noted in the discussion of STAR Early Literacy and STAR Reading validity, content is a crucial facet of test validity; content-related evidence of validity lies in the degree of correspondence, or alignment, between the knowledge and skills measured by an assessment's test items and the knowledge and skills intended to be taught and learned in a given curriculum at a given grade level or levels. STAR Math Enterprise content is aligned to curriculum standards at the state and national levels—including the Common Core State Standards (see Core Progress Learning Progressions—The Bridge Between Assessment and Instruction, p. 10).

Psychometric reliability, combined with a high degree of alignment of test content to curriculum standards may be evidence enough of an assessment's validity. However, other measures complement or corroborate those two facets and serve to further strengthen an assessment's claims of validity.

To support STAR Math Enterprise as a measure of a broad range of mathematics skills, Renaissance Learning has collected a wide range of correlations between scores on STAR Math and scores on other recognized, established measures of different aspects of mathematics achievement, such as survey achievement tests, diagnostic math measures, and state accountability tests, among others. Table 10 summarizes the results of

more than 400 concurrent and predictive validity studies conducted for STAR Math, involving a total of more than 400,000 students. The average correlations observed in these studies range from 0.55 to 0.80; correlations in that range are considered moderate to strong. Below the table is a list of state assessments that have been found to correlate well with scores on STAR Math.

Table 10: Summary of STAR Math™ Validity Studies

Grade	Predictive			Concurrent		
	Studies	Students	Average Correlation	Studies	Students	Average Correlation
1	6	11,880	.55	6	179	.58
2	10	33,076	.63	17	987	.61
3	30	52,604	.66	49	6,400	.61
4	23	55,285	.69	49	5,823	.59
5	29	39,869	.70	58	6,873	.64
6	13	27,663	.73	37	4,202	.66
7	15	18,919	.75	29	3,361	.64
8	11	12,780	.76	29	3,713	.65
9	6	2,545	.78	13	665	.57
10	6	2,236	.79	10	334	.60
11	6	1,921	.80	10	495	.68
12	2	885	.77	9	233	.68

STAR Math™ relates to several state assessments

Studies have been conducted with STAR Math and the following assessments to statistically link[6] the tests:

- ACT EXPLORE
- Alabama Reading and Mathematics Test+ (ARMT+)
- Arizona's Instrument to Measure Standards (AIMS)
- Arkansas Augmented Benchmark Examinations (AABE)
- California Standards Tests (CST)
- Colorado—Transitional Colorado Assessment Program (TCAP)
- Connecticut Mastery Test (CMT4)
- Delaware Comprehensive Assessment System (DCAS)
- Florida Comprehensive Assessment Test 2.0 (FCAT 2.0)
- Georgia's Criterion-Referenced Competency Tests (CRCT)
- Idaho Standards Achievement Test (ISAT)
- Illinois Standards Achievement Test (ISAT)
- Indiana Statewide Testing for Education Progress-Plus (ISTEP+) Assessments
- Iowa Assessment (IA)
- Kansas State Assessment Program (KSAP)
- Kentucky Performance Rating for Educational Progress (K-PREP)
- Louisiana Educational Assessment Program (LEAP) and Integrated Educational Assessment Program (iLEAP) Assessments

---

[6] Statistical linking studies are continuously underway and Performance Reports for new states are released on a regular basis. For the most up-to-date list of state assessments with statistical links to STAR Reading or STAR Math, email research@renlearn.com. Technical manuals are also available upon request and include information on the assessments that correlate with STAR Assessments.

- Maine—New England Common Assessment Program (NECAP)
- Massachusetts Comprehensive Assessment System (MCAS)
- Michigan Educational Assessment Program (MEAP)
- Minnesota Comprehensive Assessments (MCAs)
- Mississippi Curriculum Test, Second Edition (MCT2)
- Missouri Assessment Program (MAP) Grade-Level Assessments
- Montana's Criterion-Referenced Test (CRT)
- Nevada's Criterion-Referenced Test (CRT)
- New Hampshire—New England Common Assessment Program (NECAP)
- New Jersey Assessment of Skills and Knowledge (NJ ASK)
- New Mexico Standards Based Assessments (SBA)
- New York State Assessment Program (NYSTP)
- North Carolina End-of-Grade (NC EOG) Test
- Ohio Achievement Assessments (OAA)
- Oklahoma Core Curriculum Tests (OCCT)
- Pennsylvania's System of School Assessment (PSSA)
- Rhode Island—New England Common Assessment Program (NECAP)
- South Carolina Palmetto Assessment of State Standards (SCPASS)
- South Dakota State Test of Educational Progress (DSTEP)
- Tennessee Comprehensive Assessment Program (TCAP)
- Texas—State of Texas Assessments of Academic Readiness (STAAR)
- Vermont—New England Common Assessment Program (NECAP)
- Virginia Standards of Learning (SOL)
- Washington—Measurements of Student Progress (MSP)
- West Virginia Educational Standards Test 2 (WESTEST 2)
- Wisconsin Knowledge and Concepts Examination (WKCE)
- Wyoming—Proficiency Assessments for Wyoming Students (PAWS)

Purpose and Frequency

Most schools administer STAR Enterprise assessments to all students in the fall, winter, and spring for screening purposes. If educators want to establish a trend line for students (visible in reports of STAR results) to forecast proficiency on state tests or mastery of standards, they must administer an additional test in late fall. This way, after the winter screening, three data points have been established so the software can chart students' growth trajectories.

Teachers who monitor progress more closely for specific students, in an intervention or other setting, or for instructional planning, typically test more frequently. Although STAR Assessments can be administered as often as weekly, an important general guideline is to administer assessments to students only when educators are prepared to act upon the resulting data.

Response to Intervention screening and progress monitoring

Response to Intervention (RTI)—also known as a Multi-Tier System of Supports (MTSS)— is a framework for making instructional decisions based on data, in order to accelerate learning for all students. Interim assessments play a key role in RTI, helping to provide data to inform and improve instruction. Interim assessments are generally used for screening/benchmarking or progress monitoring. STAR Enterprise assessments are used for both of these purposes:

- **Screening and benchmarking** periodic assessment, typically administered two to four times per year to monitor growth of a group toward a proficiency target, which also may provide information about the standards students have likely mastered.

- **Progress-monitoring assessment**—defined as measures of academic performance by the National Center on Response to Intervention—administered more frequently than annually, but as often as monthly to monitor students' growth trajectories or weekly in intervention situations to measure individual student progress. Progress-monitoring assessments measure growth during the year and longitudinally over two or more years. Also included in this category are diagnostic assessments administered as needed to help identify specific areas of weakness. (For more information, see Computer Adaptive Testing, p. 7.)

Growth measurement: Scaled score, growth norms, and student growth percentile

Because changes in student achievement do not happen overnight, measuring growth is essential to understanding the effects of instruction. Renaissance Learning has unique insight into how students grow through ongoing study of data from the millions of tests taken by students at thousands of schools. During the 2012–2013 school year alone, more than 45 million STAR tests were taken. With this wealth of data, we are able to calculate growth norms. We can approximate how much growth is typical for students of different achievement levels in different grades from one time period to another.

In addition to screening students to forecast proficiency on end-of-year summative tests and progress monitoring their growth throughout the year, teachers can use STAR Enterprise assessments to capture a picture of each student's overall growth from the beginning of the school year to the end, or in semester increments.

> During the 2012–2013 school year alone, more than 45 million STAR tests were taken.

Scaled score

STAR Enterprise assessments generate a scaled score (SS), which is useful for comparing student performance over time. The same range is used for all students, so scaled scores help to compare student

performance across grade levels. Any scaled score increase indicates that a student has experienced growth. STAR Reading Enterprise and STAR Math Enterprise scaled scores range from 0–1400, while STAR Early Literacy Enterprise scaled scores range from 300–900 and relate directly to specific literacy classifications (Emergent Reader, Transitional Reader, and Probable Reader).

Growth norms
Just as meteorologists use statistical models to predict the weather, educational researchers use growth models to identify patterns in student growth. Renaissance Learning has developed such a model based on study of the growth patterns for millions of students. Growth norms indicate typical rates of growth per week and are differentiated by subject, grade, and starting score. These norms are updated every year, as more and more students take STAR Enterprise assessments, to ensure the growth rates reflect the most up-to-date data possible.

Currently, the STAR Reading Enterprise data set includes more than 3.5 million students, STAR Math Enterprise includes more than 2.2 million students, and STAR Early Literacy Enterprise includes more than 400,000 students. Using this information, STAR software is able to provide a projected scaled score for the end of the year, based on a growth rate achieved by 50 percent of students with a similar percentile rank as the student for whom you are setting goals. This information appears on STAR State Standards Reports and provides educators with key information about how students grow over time.

STAR growth norms also drive the Goal-Setting Wizard (see Figure 14), which helps educators set challenging, but reasonable, progress-monitoring goals personalized to each student.

Figure 14: Goal-Setting Wizard
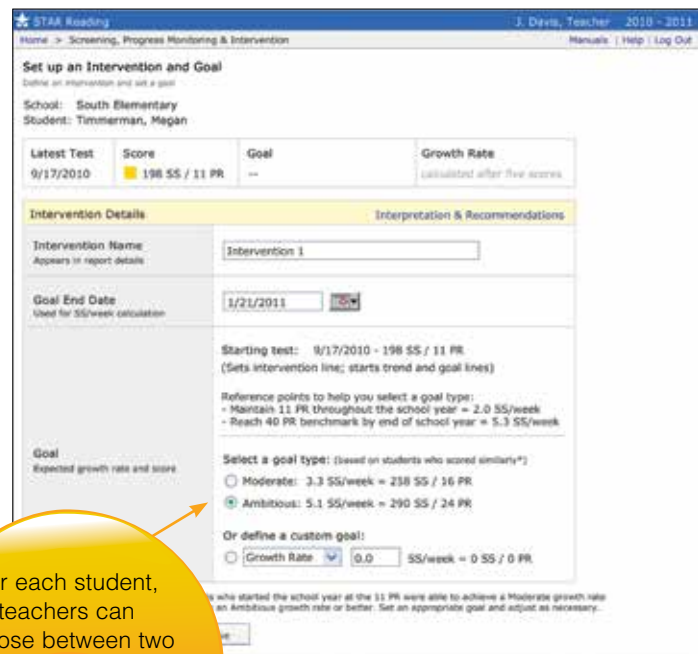
For each student, teachers can choose between two research-based recommendations for goal setting or determine a custom goal.

Student growth percentile
Student growth percentile (SGP) was first developed by Dr. Damian Betebenner from the National Center for the Improvement of Educational Assessment, in partnership with the Colorado Department of Education. Dr. Dan Bolt, at the University of Wisconsin-Madison, assisted Renaissance Learning in adapting SGP for STAR Assessments.

> SGP compares a student's growth to that of his/her *academic peers* nationwide and helps educators understand student growth.

SGP compares a student's growth to that of his/her *academic peers* nationwide and helps educators understand student growth. A student's academic peers are students at the same grade level and at similar achievement levels as that student.

An advantage of SGP is that it gives a clear picture of whether a student's growth is more or less than can be expected. A student must take at least two STAR Enterprise assessments during a school year, within specific testing windows (fall to winter, winter to spring, or fall to spring), in order to generate an SGP score and measure growth. This score helps educators at the classroom, school, and district level address important questions via tools such as reports and the Growth Proficiency Chart (see Figure 15).

For teachers:
- Did students grow from one testing period to the next?
- Did students of all abilities grow?
- Did students grow as much as expected? More? Less?
- Did students in intervention grow more than their peers nationwide?
- Did my intervention strategies lead to greater growth?

For administrators:
- How much did all students in my district grow?
- Did students of all abilities grow?
- Did students grow as much as expected? More? Less?
- Did students in some classes, grades, or schools grow more than others? What does that say about our core curriculum, intervention strategies, and programs and/or professional development needs?

Figure 15: Growth Proficiency Chart
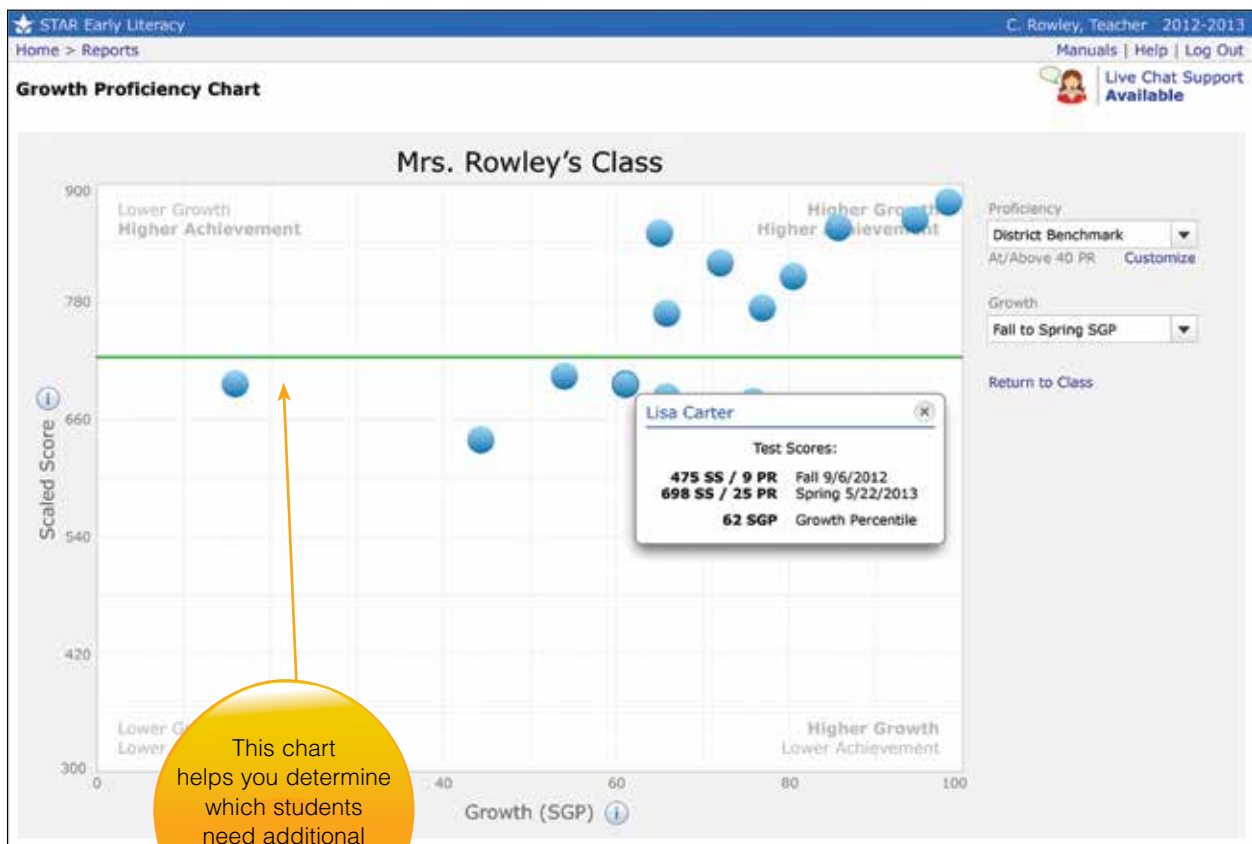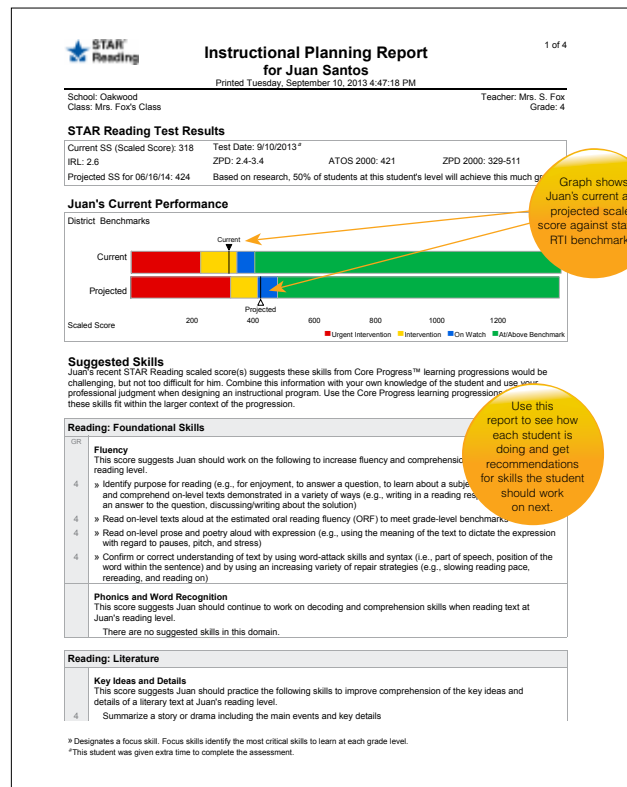
Instructional planning with Core Progress™

As mentioned, after a student takes a STAR Enterprise assessment, the software places the resulting scaled score on the Core Progress learning progression, which reports skills the student has likely mastered, those they are ready to develop next, and suggestions for the teacher to focus instruction. In essence, Core Progress serves as a road map to help teachers understand both where students have been and where they need to go to become college and career ready.

Instructional Planning Reports from STAR Enterprise provide teachers lists of skills individual students—and at the class level, groups of students—are ready to develop next (see student example, Figure 16). Within Core Progress, teachers can search for the skills and domains listed on the reports to further focus next steps for students.

Figure 17 shows a visual of the Core Progress software. Within each domain, headings match those outlined in the CCSS, and under each heading, grade-level domain expectations are identified. The software also provides resources for instruction, including Worked Examples, Skill Probes, Performance Tasks, and links to third-party educational resources.

The Record Book is another pathway to see suggested skills with which students need additional practice. This resource is especially helpful for teachers of students who need intervention, in that it suggests skills for differentiated instruction and allows teachers to create instructional groups designed for specific student needs.
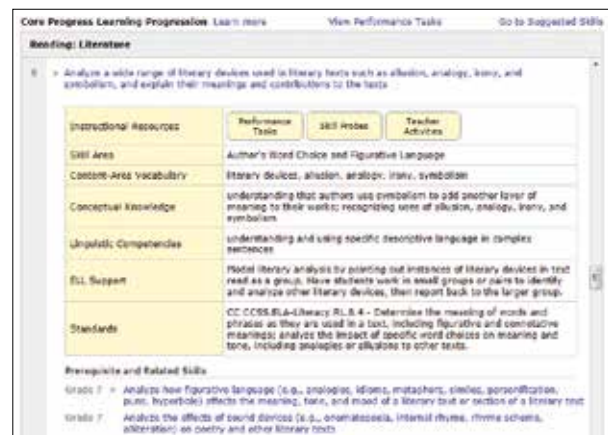
Figure 16: Instructional Planning Report

Figure 17: Core Progress™ Learning Progression for Reading—Built for the Common Core State Standards Example Screen

Predicting achievement: Linking studies and performance reporting

Will my students perform well on the state test? is one of the most serious and challenging questions teachers and administrators face. STAR Enterprise assessments are integral tools for educators to use to evaluate student progress toward proficiency.

Because STAR Assessments are computerized, achievement data for millions of students nationwide is collected each year. The Research Department at Renaissance Learning has analyzed this data and linked student performance on STAR Reading and STAR Math to student performance on several summative end-of-year state tests. (For a full list of state assessments to which STAR Reading and STAR Math have been linked, see Psychometric Properties, pp. 23, 26.)

The linking studies combined with the Renaissance Learning growth model (see Growth Norms, p. 29), which is based on STAR test results from millions of students, drive the information displayed in STAR Enterprise State Performance Reports. With versions available at the student, class, and district levels, these reports are used to monitor proficiency not only periodically, but also, more importantly, early. This way, educators know whether students are on track to achieve proficiency on the state test, and if not, they can make key instructional decisions while there is still time to do so.

Two of the reports are specifically for teachers (see Figures 18 and 19):

- **State Performance Report—Student**: Graphs a student's STAR Reading or STAR Math scores and trend line (indicates projected growth) for easy comparison with the pathway to proficiency on state reading and math tests.

- **State Performance Report—Class**: Provides a trend line at the class level depicting the average STAR Reading or STAR Math scaled score, making group progress available at a glance. Also lists individual student scores and categorizes performance as Below or On the pathway.
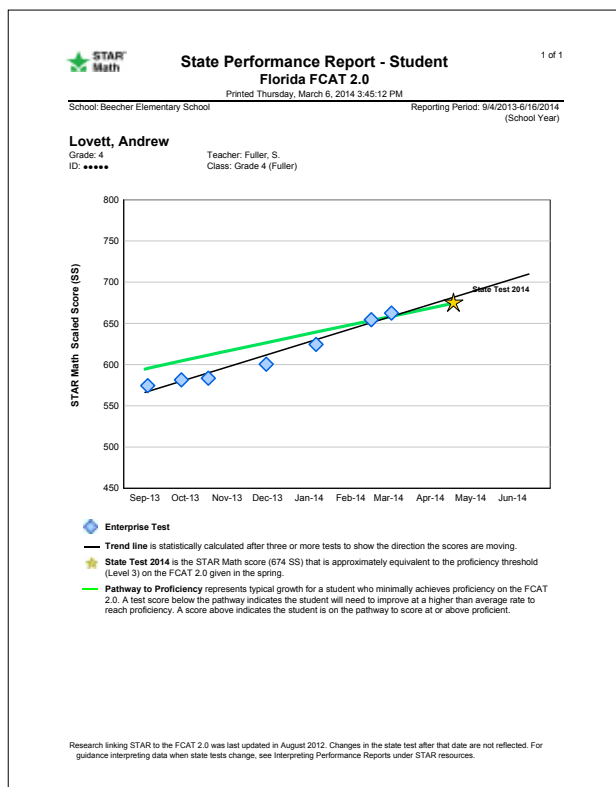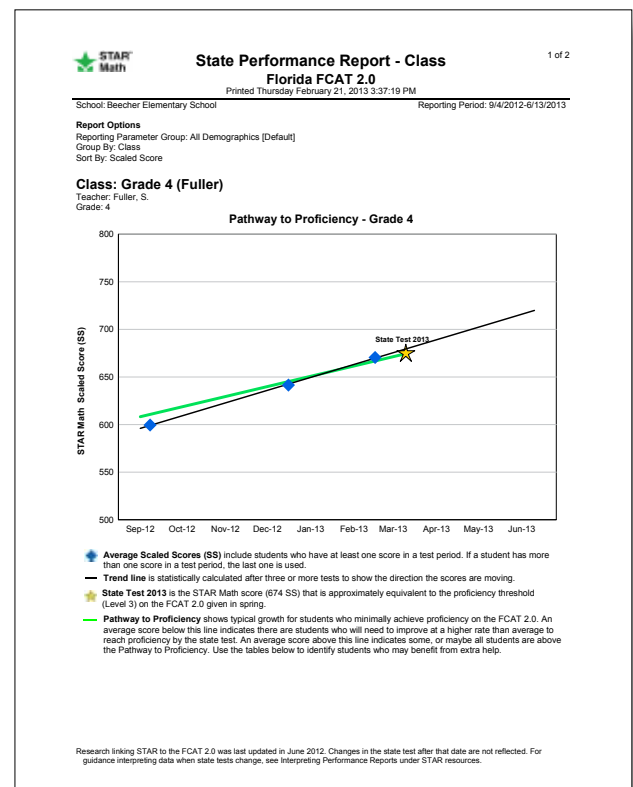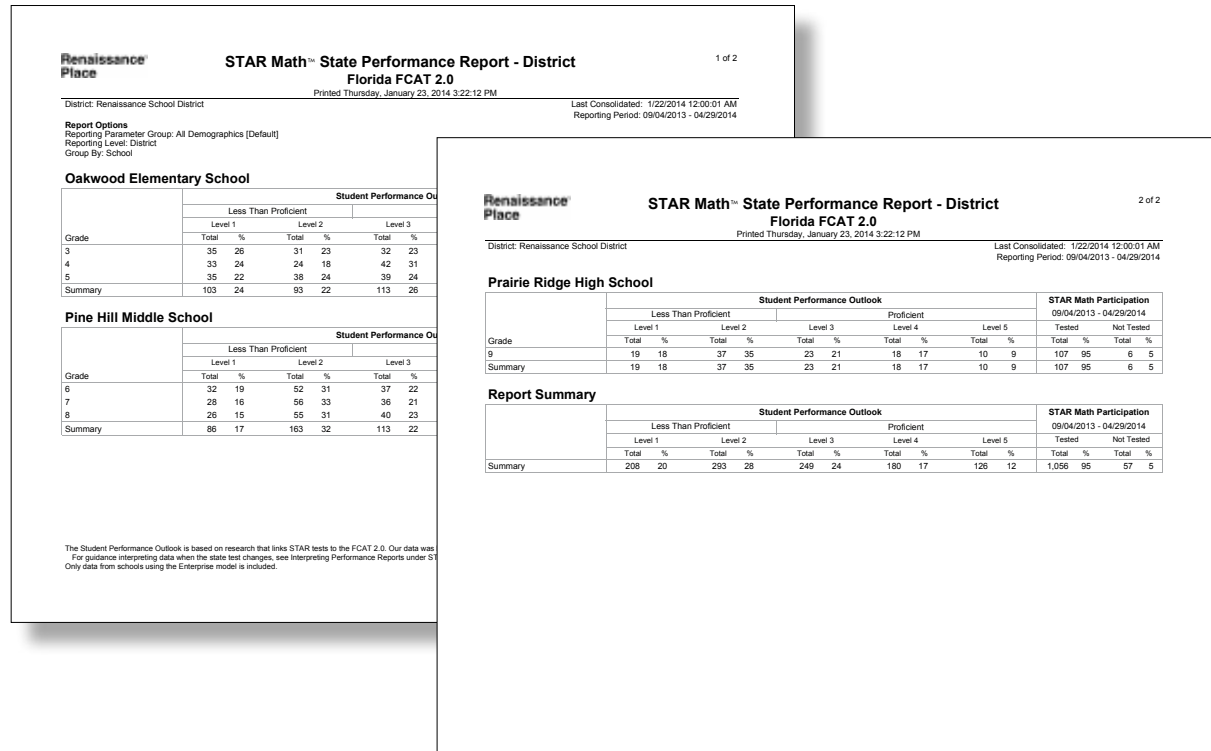
Figure 18: State Performance Report—Student

Figure 19: State Performance Report—Class

The third report is geared toward administrators (see Figure 20):

- **State Performance Report—District**: Provides a high-level performance view during the specified reporting period for each state performance level.

Figure 20: State Performance Report—District

Renaissance Place

STAR Math™ State Performance Report - District
Florida FCAT 2.0
Printed Thursday, January 23, 2014 3:22:12 PM

1 of 2

District: Renaissance School District

Last Consolidated: 1/22/2014 12:00:01 AM
Reporting Period: 09/04/2013 - 04/29/2014

Report Options
Reporting Parameter Group: All Demographics [Default]
Reporting Level: District
Group By: School

Oakwood Elementary School

	Student Performance Outlook					
	Less Than Proficient					
	Level 1		Level 2		Level 3	
Grade	Total	%	Total	%	Total	%
3	35	26	31	23	32	23
4	33	24	24	18	42	31
5	35	22	38	24	39	24
Summary	103	24	93	22	113	26

Pine Hill Middle School

	Student Performance Outlook					
	Less Than Proficient					
	Level 1		Level 2		Level 3	
Grade	Total	%	Total	%	Total	%
6	32	19	52	31	37	22
7	28	16	56	33	36	21
8	26	15	55	31	40	23
Summary	86	17	163	32	113	22

The Student Performance Outlook is based on research that links STAR tests to the FCAT 2.0. Our data was...
For guidance interpreting data when the state test changes, see Interpreting Performance Reports under ST...
Only data from schools using the Enterprise model is included.

Renaissance Place

STAR Math™ State Performance Report - District
Florida FCAT 2.0
Printed Thursday, January 23, 2014 3:22:12 PM

2 of 2

District: Renaissance School District

Last Consolidated: 1/22/2014 12:00:01 AM
Reporting Period: 09/04/2013 - 04/29/2014

Prairie Ridge High School

	Student Performance Outlook										STAR Math Participation			
	Less Than Proficient				Proficient						09/04/2013 - 04/29/2014			
	Level 1		Level 2		Level 3		Level 4		Level 5		Tested		Not Tested	
Grade	Total	%	Total	%	Total	%	Total	%	Total	%	Total	%	Total	%
9	19	18	37	35	23	21	18	17	10	9	107	95	6	5
Summary	19	18	37	35	23	21	18	17	10	9	107	95	6	5

Report Summary

	Student Performance Outlook										STAR Math Participation			
	Less Than Proficient				Proficient						09/04/2013 - 04/29/2014			
	Level 1		Level 2		Level 3		Level 4		Level 5		Tested		Not Tested	
	Total	%	Total	%	Total	%	Total	%	Total	%	Total	%	Total	%
Summary	208	20	293	28	249	24	180	17	126	12	1,056	95	57	5

Standards alignment and reporting with the Common Core and other state standards

The Renaissance Learning standards team actively follows best practices in standards research and alignment, as well as maintains ongoing relationships in research and consultation with leading educational organizations, such as Mid-continent Research for Education and Learning (McREL) and the Northwest Regional Educational Laboratory (NWREL). This team rigorously developed, tested, and validated the original Core Progress learning progressions, and in 2013, they fully immersed themselves in the Common Core State Standards, as well as literature, resources, interpretations, and implementation information surrounding them to create another set of learning progressions built specifically for the Common Core.

STAR State Standards Reports (see example, Figure 21, next page), generated by the STAR Enterprise software, help educators estimate a student, class, or district's level of mastery on the Common Core State Standards or individual state standards (for those states that have not adopted the CCSS). To develop these reports, the standards team used both empirical data and content-area expert review, similar to the method used by states to place their standards on state test scales. Standards were aligned with the STAR scale using the following rigorous five-step approach:

— **33** —

1. Identify standards

2. Identify STAR skills and the items for those skills that assess the standard's expectations for the skills and concepts.

3. Review the calibrated (research-based) difficulty level of STAR items associated with the skills and concepts embedded in the standard.

4. Assign a difficulty level to the standard based on the review of empirical data. Equate the difficulty level to a scaled score on the STAR scale.

5. Review of assigned STAR scaled score by a content-area expert who analyzes the assigned score in relation to the composite standard to ensure the placement is accurate and appropriate.

High stakes purposes

Educators use assessments for different purposes. Some assessments can only be used for a single purpose, while others, such as STAR Enterprise, can meet various needs. Many of the uses of STAR described in this document are instructional—helping teachers understand what students know and what they are ready to learn next, how much they are growing, or whether they are responding adequately to instruction. Yet as educators well know, states and districts have been using assessment results for other, higher stakes decisions. STAR Enterprise assessments are approved by many states and districts for such purposes, typically as one of multiple measures or data points. These purposes include, but are not limited to, the following:

- Serving as an indicator of student growth in educator evaluation formulas
- Grade promotion
- Gifted & Talented identification

Uses of STAR Assessments for these purposes depend on specific state and district policies, but one commonality among them is that they demand assessments show evidence of strong technical adequacy, including reliability, validity, and predictive accuracy. The fact that STAR Assessments are often approved for these uses provides further reinforcement that the assessments meet high technical standards.

Figure 21: State Standards Report—Student (Common Core State Standards)

34
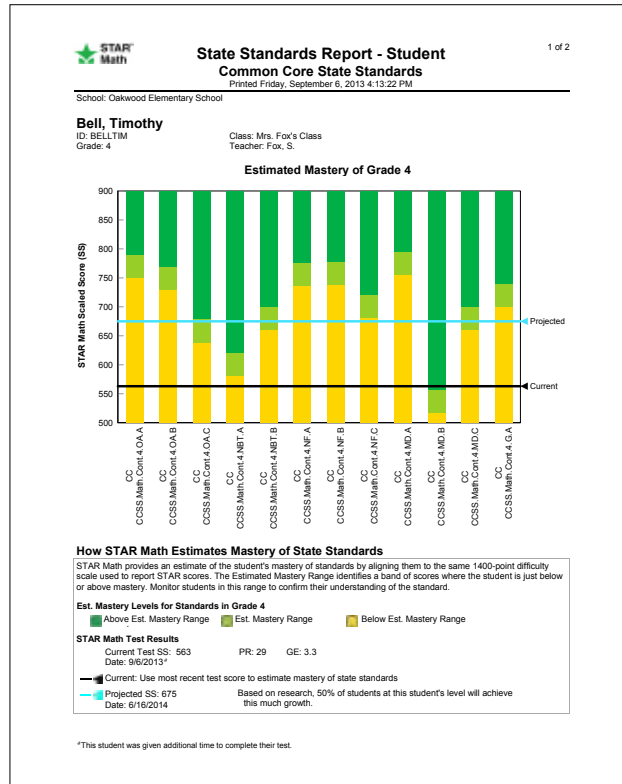
Test Content

Large item banks

STAR Assessments have large item banks to allow multiple administrations without risk of item overexposure. The STAR Early Literacy Enterprise item bank contains more than 2,500 items, while the STAR Reading Enterprise and STAR Math Enterprise items banks each number more than 5,000 items. Renaissance Learning continually develops new high-quality assessment items that are added to the banks to support frequent testing and to achieve an even distribution of items across the difficulty levels of each STAR assessment.

STAR Enterprise assessments are fixed-length tests, which mean item count is the sole criterion for ending an administration. STAR Early Literacy Enterprise administers 27 items per test event, and STAR Reading Enterprise and STAR Math Enterprise each administer 34 items. The tests were developed to provide precise measurement of student achievement in early literacy (and early numeracy), reading, and math, and to do so efficiently. Because the assessments are computer adaptive tests (CATs), they save teachers time by automating administration and scoring. Even more importantly, CATs allow students to be assessed on a larger and more varied range of skills using fewer items, which results in students spending less time completing the assessment (for more information, see Test Design, p. 7).

> CATs allow students to be assessed on a larger and more varied range of skills using fewer items, which results in students spending less time completing the assessment.

Multiple-choice format

Renaissance Learning examined, researched, discussed, and prototyped several item-response formats and ultimately chose to use multiple-choice test items. Much research supports the use of this item type, also referred to as selected-response format. As noted by Stiggins (2005):

> [Selected-response] tests are efficient in that we can administer large numbers of multiple-choice or true/false test items per unit of testing time. Thus, they permit us to sample widely and draw relatively confident generalizations from the content sampled. For this reason, when the target is knowledge mastery, selected-response formats fit nicely into the resource realities of most classrooms. (p. 70)

The multiple-choice format lends itself well to computerized scoring, which automates the testing process and saves teachers time in collecting and scoring results (Nicol, 2007). A large number of multiple-choice test items can be administered in a short amount of time, and a key factor in the measurement precision of any test is the number of items each student must answer. According to Haladyna and Downing (1989), "the use of multiple-choice formats generally leads to more content-valid test score interpretations."

Renaissance Learning constructs multiple-choice items to represent a balanced range of cognitive complexity. Item specifications require verifying the accuracy of all content; using grade-level-appropriate cognitive load, vocabulary, syntax, and readability; including only essential text and graphics to avoid wordiness and visual clutter; and employing standards for bias, fairness, and sensitivity.

Research has shown that well-designed multiple-choice questions can assess an array of skills (Cassels & Johnstone, 1984; Popham, 2008; Russell, Fischer, Fischer, & Premo, 2003) at higher levels of student learning (Cox, 1976; Johnstone & Arnbusaidi, 2000; Mattimore, 2009; Osterlind, 1998; Popham, 2003).

Item-development process

Item development is of critical concern to Renaissance Learning. The care in developing items is reflected in the high ratings STAR Assessments have garnered from several key federal groups, such as the National Center on Intensive Intervention, the National Center on Response to Intervention, and then National Center on Student Progress Monitoring (for more information, see Psychometric Properties, p. 19).

Professional designers, writers, and editors—with education backgrounds and content-area expertise—develop all content for Renaissance Learning products, including STAR Enterprise assessments. These experts follow research-based practices for developing assessment items, and rigorously adhere to the following process to ensure quality item creation:

1. Analyze standards to be assessed in the categories of skill, action, vocabulary, and context; refer to national or state resources for appropriate standard and grade-level expectation interpretation.
2. Write item specifications and provide training on their use to item writers and editors.
3. Establish item metadata to guide development, including standards-related and item-related data.
4. Use a multistep, recursive writing and editing process that ensures adherence to specifications and alignment to standards and item metadata.
5. Post items for calibration and acquire student-response data dynamic calibration (see below).
6. Examine psychometricians' analyses of item-testing results.
7. Add successful items to the operational assessment item bank.

Experts also receive ongoing item-writing training, which includes bias-and-fairness criteria to avoid stereotypes and characterizations of people or events that could be construed as demeaning, patronizing, or otherwise insensitive. Content-development tools track and report attributes such as gender, age, ethnicity, subject matter, and regional references. Individual attributes, as well as the intersection of multiple attributes, are tracked throughout the development process to ensure that final content is demographically balanced and free of bias.

In addition, assessment items must also pass strict quality reviews which check for discipline-specific criteria, accuracy, language appropriateness and readability level, bias and fairness, and technical quality control.

Rules for item retention

Following these steps, all information pertaining to each test item—including traditional- and IRT-analysis data, test level, form, and item identifier—is stored in an item-statistics database. Then a panel of content reviewers examines each item within content strands to determine whether the item meets all criteria for use in an operational assessment. After all content reviewers have designated any items for elimination, the recommendations are combined and a second review is conducted to resolve any issues.

Dynamic calibration

To maintain and update the large item banks for each STAR assessment, Renaissance Learning continually develops and calibrates new test items using a special feature called dynamic calibration. Each new STAR assessment item goes through calibration to determine its exact point on the STAR difficulty scale.

In dynamic calibration, one or more new items are embedded at random points in a STAR test. The items are administered to large samples of students, so that Renaissance Learning psychometricians can collect student-response and other data on the item, and then perform a statistical analysis of the response data to determine the scale values.

These items do not count toward students' scores on the STAR assessment. Students, on average, receive two or three additional items per test when calibration is turned on, and testing time is increased by approximately one minute. Norming, reliability, and validity studies take place after items successfully pass through calibration.

Appendix: STAR Assessments™ Score Definitions

STAR Early Literacy Enterprise™ scores

Literacy Classifications are the stages of literacy development measured in STAR Early Literacy and associated with scaled scores. They are an easy way to monitor student progress:

> **Emergent Reader (300–674)**:
> An *Early Emergent Reader (300–487)* is beginning to understand that printed text has meaning. The student is learning that reading involves printed words and sentences and that print flows from left to right and from top to bottom of a page. The student is also beginning to identify colors, shapes, numbers, and letters.
>
> A *Late Emergent Reader (488–674)* can identify most of the letters of the alphabet and match most of the letters to sounds. The student is beginning to "read" picture books and familiar words around home. Through repeated reading of favorite books with an adult, a student at this stage is building vocabulary, listening skills, and understanding of print.
>
> A **Transitional Reader (675–774)** has mastered alphabet skills and letter-sound relationships. The student can identify many beginning and ending consonant sounds as well as long and short vowel sounds. The student is probably able to blend sounds and word parts to read simple words and is likely using a variety of strategies to figure out words, such as pictures, story patterns, and phonics.
>
> A **Probable Reader (775–900)** is becoming proficient at recognizing many words, both in and out of context, and spends less time identifying and sounding out words and more time understanding what was read. A probable reader can blend sounds and word parts to read words and sentences more quickly, smoothly, and independently than students in other stages of development.

Literacy Domain Score, ranging from 0–100, is criterion-referenced and represents the percentage of items a student would be expected to answer correctly within the assessment's domains, which include key early literacy sub-domains comprised of skill sets.

Sub-Domain and **Skill Set Scores**, ranging from 0–100, are criterion-referenced and represent the percent of mastery of specific skills within the assessment's domains, sub-domains, and skill sets.

Estimated Oral Reading Fluency (Est. ORF), reported in correct words per minute, is an estimate of a student's ability to read words quickly and accurately in order to comprehend text efficiently. Students with oral reading fluency demonstrate accurate decoding, automatic word recognition, and appropriate use of the rhythmic aspects of language (e.g., intonation, phrasing, pitch, emphasis). Est. ORF is based on a known relationship between STAR Early Literacy Enterprise performance and oral reading fluency and is reported for grades 1–4.

Growth Norms characterize typical student growth within a given grade and achievement level.

Scaled Score (SS) is useful in comparing student performance over time and in identifying student performance in relation to a vertical scale and all criteria associated with that scale. Because the same range is used for all students, scaled scores are also useful for comparing student performance across grade levels. STAR Early Literacy Enterprise scaled scores range from 300–900.

Student Growth Percentile (SGP) is a measure of growth between a pre- and posttest, relative to the growth made by other students in the same grade with the same pretest score. It is a simple and effective way for educators to interpret student growth rate relative to that of his or her academic peers nationwide. SGPs, which were derived from growth norms, range from 1–99, with lower numbers representing lower relative growth and high numbers representing higher relative growth.

STAR Reading Enterprise™ scores
ATOS 2000 is the STAR scaled score converted to Renaissance Learning's 2000-point scale, based on an extensive research study correlating STAR to the Lexile scale. While it is not a Lexile score, it is intended to provide a score that can be used in place of a Lexile score as a close approximation.

Domain and **Skill Set Scores**, ranging from 0–100, are criterion-referenced and estimate a student's percent of mastery of specific skills within the assessment's domains and skill sets.

Estimated Oral Reading Fluency (Est. ORF), reported in correct words per minute, is an estimate of a student's ability to read words quickly and accurately in order to comprehend text efficiently. Students with oral reading fluency demonstrate accurate decoding, automatic word recognition, and appropriate use of the rhythmic aspects of language (e.g., intonation, phrasing, pitch, emphasis). Est. ORF is based on a known relationship between STAR Reading Enterprise performance and oral reading fluency and is reported for grades 1–4.

Grade Equivalent (GE) score, ranging 0.0–12.9+, is norm-referenced and represents how a student's test performance compares with other students nationally. For example, a fifth-grade student with a GE of 7.6 performed as well as a typical seventh-grader in the sixth month of the school year. This does not mean the student is necessarily capable of reading seventh-grade material—rather, it indicates that the student's reading skills are well above average for fifth grade.

Growth Norms characterize typical student growth within a given grade and achievement level.

Instructional Reading Level (IRL) is a criterion-referenced score that is the highest reading level at which a student is 80% proficient (or higher) at comprehending material with assistance (Gickling & Thompson, 2001). Research has found that this level of comprehension corresponds to being at least 90–98% proficient at recognizing words (Gickling & Havertape, 1981; Johnson, Kress, & Pikulski, 1987; McCormick, 1999).[7] IRL scores are PP (Pre-Primer), P (Primer, grades 0.1–0.9), grades 1.0 through 12.9, and PHS (Post-High School, grades 13.0+).

Normal Curve Equivalent (NCE) score, ranging from 1–99, is norm-referenced and similar to the percentile rank score but based on an equal interval scale. This means the difference between any two successive scores on the NCE scale has the same meaning throughout the scale. Mostly used for research, NCEs are useful in making comparisons between different achievement tests and in statistical computations—for example, determining an average score for a group of students.

Percentile Rank (PR) score, ranging from 1–99, is norm-referenced and provides the best measure of a student's reading achievement level compared to other students in the same grade nationally. The score indicates the percentage of a student's peers whose scores were equal to or lower than the score of that student—for example, a student with a PR score of 85 performed as well as or better than 85 percent of students in the same grade.

Scaled Score (SS) is useful in comparing student performance over time and in identifying student performance in relation to a vertical scale and all criteria and norms associated with that scale. Because the same range is used for all students, scaled scores are also useful for comparing student performance across grade levels. STAR Reading Enterprise scaled scores range from 0–1400.

Student Growth Percentile (SGP) is a measure of growth between a pre- and posttest relative to the growth made by other students in the same grade with the same pretest score. It is a simple and effective way for educators to interpret student growth rate relative to that of his or her academic peers nationwide. SGPs, which were derived from growth norms, range from 1–99, with lower numbers representing lower relative growth and high numbers representing higher relative growth.

Zone of Proximal Development (ZPD) is an individualized range of readability levels based on a student's results from a STAR Reading Enterprise assessment. Books students choose to read within their ZPD range will be neither too difficult nor too easy and should allow students to experience optimal growth.

STAR Math Enterprise™ scores
Accelerated Math Library Recommendation helps educators place a student in the Accelerated Math library that will be of the most benefit, based on that student's achievement level per the results of a STAR Math Enterprise assessment.

Algebra Readiness Indicator is based solely on skills associated with algebra readiness. The math concepts and skills learned in elementary through middle school provide the foundation for high school level algebra. The Student Instructional Planning Report in STAR Math Enterprise provides an Algebra Readiness Indicator to help teachers identify student progress through these foundational skills to ensure the student is on track to be ready for algebra.

Domain and **Skill Set Scores**, ranging from 0–100, are criterion-referenced and estimate a student's percentage of mastery of specific skills within the assessment's domains and skill sets.

Grade Equivalent (GE) score, ranging 0.0–12.9+, is norm-referenced and represents how a student's test performance compares with other students nationally. For example, a fifth-grade student with a GE of 7.6 performed as well as a typical seventh-grader in the sixth month of the school year. This does not mean the student is necessarily capable of doing seventh-grade math—rather, it indicates that the student's math skills are well above average for fifth grade.

Growth Norms characterize typical student growth within a given grade and achievement level.

Normal Curve Equivalent (NCE) score, ranging from 1–99, is norm-referenced and similar to the percentile rank score but based on an equal interval scale. This means the difference between any two successive scores on the NCE scale has the same meaning throughout the scale. Mostly used for research, NCEs are useful in making comparisons between different achievement tests and in statistical computations—for example, determining an average score for a group of students.

[7] Gickling, E. E., & Havertape, S. (1981). *Curriculum-based assessment (CBA).* Minneapolis, MN: School Psychology Inservice Training Network.

Gickling, E. E., & Thompson, V. E. (2001). Putting the learning needs of children first. In B. Sornson (Ed.). *Preventing early learning failure.* Alexandria, VA: ASCD.

Johnson, M. S., Kress, R. A., & Pikulski, J. J. (1987). *Informal reading inventories.* Newark, DE: International Reading Association.

McCormick, S. (1999). *Instructing students who have literacy problems* (3rd Ed.). Englewood Cliffs, NJ: Prentice-Hall.

Percentile Rank (PR) score, ranging from 1–99, is norm-referenced and provides the best measure of a student's math achievement level compared to other students in the same grade nationally. The score indicates the percentage of a student's peers whose scores were equal to or lower than the score of that student—for example, a student with a PR score of 85 performed as well as or better than 85 percent of students in the same grade.

Scaled Score (SS) is useful in comparing student performance over time and in identifying student performance in relation to a vertical scale and all criteria and norms associated with that scale. Because the same range is used for all students, scaled scores are also useful for comparing student performance across grade levels. STAR Math Enterprise scaled scores range from 0–1400.

Student Growth Percentile (SGP) is a measure of growth between a pre- and posttest relative to the growth made by other students in the same grade with the same pretest score. It is a simple and effective way for educators to interpret student growth rate relative to that of his or her academic peers nationwide. SGPs, which were derived from growth norms, range from 1–99, with lower numbers representing lower relative growth and high numbers representing higher relative growth.

References

Bray, M. A. & Kehle, T. J., (Eds.) (2011). *The Oxford handbook of school psychology.* New York: Oxford University Press, Inc.

Cassels, J. R. T., & Johnstone, A. H. (1984). The effect of language on student performance on multiple choice tests in chemistry. *Journal of Chemistry Education, 61*, 613–615.

Cox, K. R. (1976). How did you guess? Or what do multiple choice questions measure? *Medical Journal of Australia, 1*, 884–886.

Haladyna, T. M., & Downing, S. M. (1989). The validity of a taxonomy of multiple-choice item-writing rules. *Applied Measurement in Education, 1*, 51–78.

Heritage, M. (2008). *Learning progressions: Supporting instruction and formative assessment.* Washington, DC: Council of Chief School Officers. Retrieved from http://www.ccsso.org/Documents/2008/Learning_Progressions_Supporting_2008.pdf

Johnson, M. J., & Weiss, D. J. (1980). Parallel forms reliability and measurement accuracy comparison of adaptive and conventional testing strategies. In D. J. Weiss (Ed.), *Proceedings of the 1979 Computerized Adaptive Testing Conference* (pp. 16–34). Minneapolis: University of Minnesota, Department of Psychology, Psychometric Methods Program, Computerized Adaptive Testing Laboratory. Retrieved from http://www.psych.umn.edu/psylabs/catcentral/pdf%20files/jo80-01.pdf

Johnstone, A. H., & Ambusaidi, A. (2000). Fixed response: What are we testing? *Chemistry Education: Research and Practice in Europe, 1*(3), 323–328.

Kingsbury, G. G., & Weiss, D. J. (1980). *An alternate-forms reliability and concurrent validity comparison of Bayesian adaptive and conventional ability tests* (Research Report 80-5). Minneapolis: University of Minnesota, Department of Psychology, Psychometric Methods Program, Computerized Adaptive Testing Laboratory. Retrieved from http://iacat.org/sites/default/files/biblio/ki80-05.pdf

Lord, F. M. (1980). *Applications of item response theory to practical testing problems* (pp. 158–159). Hillsdale, NJ: Lawrence Erlbaum Associates.

Mardberg, B., & Carlstedt, B. (1998). Swedish Enlistment Battery: Construct validity and latent variable estimation of cognitive abilities by the CAT-SEB. *International Journal of Selection and Assessment, 6*, 107–114. Retrieved from http://onlinelibrary.wiley.com/doi/10.1111/1468-2389.00079/abstract

Mattimore, P. (2009, February 5). *Why our children need national multiple choice tests.* Retrieved from http://www.opednews.com/articles/Why-Our-Children-Need-Nati-by-Patrick-Mattimore-090205-402.html

McBride, J., & Martin, J. T. (1983). Reliability and validity of adaptive ability tests. In D. J. Weiss (Ed.), *New horizons in testing: Latent trait test theory and computerized adaptive testing* (Chapter 11, pp. 224–225). New York: Academic Press.

Moreno, K. E., & Segall, O. D. (1997). Reliability and construct validity of CAT-ASVAB. In W. A. Sands, B. K. Waters, & J. R. McBride (Eds.), *Computerized adaptive testing: From inquiry to operation* (pp. 169–179). Washington, DC: American Psychological Association. Retrieved from http://psycnet.apa.org/index.cfm?fa=buy.optionToBuy&id=1997-36257-018

Nicol, D. (2007). E-assessment by design: Using multiple-choice tests to good effect. *Journal of Further and Higher Education, 31*(1), 53–64.

Osterlind, S. J. (1998). *Constructing test items: Multiple-choice, constructed-response, performance, and other formats* (2nd ed.). New York: Kluwer.

Popham, W. J. (2003). *Test better, teach better: The instructional role of assessment.* Alexandria, VA: Association for Supervision and Curriculum Development.

Popham, W. J. (2008). *Classroom assessment: What teachers need to know* (5th Ed.). Boston: Allyn and Bacon.

Renaissance Learning. (2013a). *STAR Early Literacy Enterprise: Technical manual.* Wisconsin Rapids, WI: Author. Available by request to research@renlearn.com

Renaissance Learning. (2013b). *STAR Math: Technical manual.* Wisconsin Rapids, WI: Author. Available by request to research@renlearn.com

Renaissance Learning. (2013c). *STAR Reading: Technical manual.* Wisconsin Rapids, WI: Author. Available by request to research@renlearn.com

Russell, M., Fischer, M. J., Fischer, C. M., & Premo, K. (2003). Exam question sequencing effects on marketing and management sciences student performance. *Journal of Advancement of Marketing Education, 3*, 1–10.

Shapiro, E. S. (2012). *New thinking in response to intervention: A comparison of computer-adaptive tests and curriculum-based measurement within RTI.* Wisconsin Rapids, WI: Renaissance Learning, Inc.

Shapiro, E. S., & Gebhardt, S. N. (2012). Comparing computer-adaptive and curriculum-based measurement methods of assessment. *School Psychology Review, 41*(3), 295-305.

Stiggins, R. J. (2005). *Student-involved classroom assessment for learning* (4th ed.). Upper Saddle River, New Jersey: Pearson/Merrill Prentice Hall.

U.S. Department of Education: National Center on Intensive Intervention. (2012a). *Review of progress-monitoring tools* [Review of STAR Early Literacy]. Washington, DC: Author. Available online from http://www.intensiveintervention.org/chart/progress-monitoring

U.S. Department of Education: National Center on Intensive Intervention. (2012b). *Review of progress-monitoring tools* [Review of STAR Math]. Washington, DC: Author. Available online from http://www.intensiveintervention.org/chart/progress-monitoring

U.S. Department of Education: National Center on Intensive Intervention. (2012c). *Review of progress-monitoring tools* [Review of STAR Reading]. Washington, DC: Author. Available online from http://www.intensiveintervention.org/chart/progress-monitoring

U.S. Department of Education: National Center on Response to Intervention. (2010a). *Review of progress-monitoring tools* [Review of STAR Early Literacy]. Washington, DC: Author. Available online from http://www.rti4success.org/progressMonitoringTools

U.S. Department of Education: National Center on Response to Intervention. (2010b). *Review of progress-monitoring tools* [Review of STAR Math]. Washington, DC: Author. Available online from http://www.rti4success.org/progressMonitoringTools

U.S. Department of Education: National Center on Response to Intervention. (2010c). *Review of progress-monitoring tools* [Review of STAR Reading]. Washington, DC: Author. Available online from http://www.rti4success.org/progressMonitoringTools

U.S. Department of Education: National Center on Response to Intervention. (2011a). *Review of screening tools* [Review of STAR Early Literacy]. Washington, DC: Author. Available online from http://www.rti4success.org/screeningTools

U.S. Department of Education: National Center on Response to Intervention. (2011b). *Review of screening tools* [Review of STAR Math]. Washington, DC: Author. Available online from http://www.rti4success.org/screeningTools

U.S. Department of Education: National Center on Response to Intervention. (2011c). *Review of screening tools* [Review of STAR Reading]. Washington, DC: Author. Available online from http://www.rti4success.org/screeningTools

U.S. Department of Education: National Center on Student Progress Monitoring. (2006a). *Review of progress monitoring tools* [Review of STAR Early Literacy]. Washington, DC: Author. Available online from http://doc.renlearn.com/kmnet/R004100821GG6E03.pdf

U.S. Department of Education: National Center on Student Progress Monitoring. (2006b). *Review of progress monitoring tools* [Review of STAR Math]. Washington, DC: Author. Available online from http://doc.renlearn.com/kmnet/R004100821GG6E03.pdf

U.S. Department of Education: National Center on Student Progress Monitoring. (2006c). *Review of progress monitoring tools* [Review of STAR Reading]. Washington, DC: Author. Available online from http://doc.renlearn.com/kmnet/R004100821GG6E03.pdf

Weiss, D. J. (2004). Computerized adaptive testing for effective and efficient measurement in counseling and education. *Measurement and Evaluation in Counseling and Development, 37*, 70–84. Retrieved from http://www.psych.umn.edu/psylabs/catcentral/pdf%20files/we04070.pdf

Acknowledgements

The following experts have advised Renaissance Learning in the development of STAR Assessments.

R. James Milgram, Ph.D., is a professor of mathematics at Stanford University. His work in mathematics education includes consulting with several states on math standards, including California. Milgram has given lectures around the world and is a member of numerous boards and committees, including the National Board of Education Sciences, created by the Education Sciences Reform Act of 2002 "to advise and consult with the Director of the Institute of Education Sciences (IES) on agency policies," and the Human Capital Committee of the NASA Advisory Council, which "provides the NASA Administrator with counsel and advice on programs and issues of importance to the Agency." Milgram is author of "An Evaluation of CMP," "A Preliminary Analysis of SAT-I Mathematics Data for IMP Schools in California," and "Outcomes Analysis for Core Plus Students at Andover High School: One Year Later." Each of these papers identifies serious shortcomings in popular mathematics programs.

Michael Milone, Ph.D., is a research psychologist and an award-winning educational writer and consultant to publishers and school districts. He earned a Ph.D. in 1978 from The Ohio State University and has served in an adjunct capacity at Ohio State, the University of Arizona, Gallaudet University, and New Mexico State University. He has taught in regular and special education programs at all levels, holds a Master of Arts degree from Gallaudet University, and is fluent in American Sign Language. Milone served on the board of directors of the Association of Educational Publishers and was a member of the Literacy Assessment Committee and a past chair of the Technology and Literacy Committee of the International Reading Association. He has contributed to both readingonline.org and *Technology & Learning* magazine on a regular basis. Over the past 30 years, he has been involved in a broad range of publishing projects, including the SRA reading series, assessments developed for Academic Therapy Publications, and software published by The Learning Company and LeapFrog. He has completed 34 marathons and 2 Ironman races.

James B. Olsen, Ph.D., is a psychometrician for Renaissance Learning. He has professional skills in educational and professional assessment, job analysis, test design and development, psychometric and statistical analysis, research and evaluation, and professional certification and accreditation. He serves on the Technical Advisory Committees (TAC) for the Graduate Management Admissions Council, the Comprehensive Adult Student Assessment System and the Utah State Office of Education. He has served as a consultant for the National Assessment Governing Board, National Assessment of Educational Progress, and American National Standards Institute. He was principal investigator and program manager for multiple assessment projects with the U.S. Department of Defense, National Science Foundation, and U.S. Department of Education. For the Association of Test Publishers (ATP), he has served for two terms on the Board of Directors and for five years as Innovations Chair for the annual Innovations in Testing Conference.

Sharif M. Shakrani, Ph.D., is a private consultant and researcher specializing in measurement and quantitative methods. Shakrani is a former co-director of the Education Policy Center at Michigan State University and professor of measurement and quantitative methods in the Department of Counseling, Educational Psychology and Special Education. Before coming to Michigan State University, Shakrani served 8 years as the deputy executive director of the National Assessment Governing Board in the U.S. Department of Education. He was responsible for technical and policy direction for the National Assessment of Educational Programs (NAEP). He has also worked for the National Center for Education Statistics in the U.S. Department of Education where he guided the design and analysis of federal educational assessments. In his work in the Michigan Department of Education, Shakrani was responsible for K–12 general curriculum and assessment and was instrumental in revising the Michigan Educational Assessment Program (MEAP).

Amanda M. VanDerHeyden, Ph.D., is a private consultant and researcher, living in Fairhope, Alabama, who has previously held faculty positions and has worked as a researcher, consultant, and national trainer in a number of school districts. In 2006, VanDerHeyden was named to a National Center for Learning Disabilities advisory panel to provide guidance related to RTI and the diagnosis of specific learning disability. She is associate editor of *Assessment for Effective Intervention*, serves on the editorial boards of several journals including *School Psychology Quarterly* and *Journal of School Psychology*, and has recently co-authored *Essentials of Response to Intervention* (with Dr. Matthew Burns). VanDerHeyden received the 2006 Lightner Witmer Early Career Contributions Award from the APA for her scholarship on early intervention, RTI, and models of data-based decision making. She serves as research advisor to iSTEEP, has published measures of early numeracy for preschool and kindergarten children, and currently directs research projects in early numeracy with preschool/kindergarten children as well as evaluation of Tier 2 mathematics intervention.

James Ysseldyke, Ph.D., is Emma Birkmaier Professor of Educational Leadership in the Department of Educational Psychology at the University of Minnesota. He has been educating school psychologists and researchers for more than 35 years. Ysseldyke has served the University of Minnesota as director of the Minnesota Institute for Research on Learning Disabilities, director of the National School Psychology Network, director of the National Center on Educational Outcomes, director of the School Psychology Program, and associate dean for research. His research and writing have focused on enhancing the competence of individual students and enhancing the capacity of systems to meet students' needs. He is an author of major textbooks and more than 300 journal articles. Ysseldyke is conducting a set of investigations on the use of technology-enhanced progress-monitoring systems to track the performance and progress of students in urban environments. He chaired the task forces that produced the three *Blueprints on the Future of Training and Practice in School Psychology*, and he is former editor of *Exceptional Children*, the flagship journal of the Council for Exceptional Children. Ysseldyke has received awards for his research from the School Psychology Division of the American Psychological Association, the American Educational Research Association, and the Council for Exceptional Children. The University of Minnesota presented him a distinguished teaching award, and he received a distinguished alumni award from the University of Illinois.

Acknowledgements