The STAR assessments are highly rated for screening and progress monitoring by the U.S. Department of Education's **National Center on Response to Intervention**!

THE FOUNDATION OF THE

STAR Assessments

The Science of STAR

Renaissance Learning™

Advanced Technology for Data-Driven Schools™

Quick-Reference Guide to the STAR Assessments

STAR Early Literacy—used for screening, progress-monitoring, and diagnostic assessment—is a reliable, valid, and efficient, computer-adaptive assessment of 41 skills in seven critical early literacy domains. A STAR Early Literacy assessment can be completed without teacher assistance in about 10 minutes by emergent readers in grades pre-K–3 and repeated as often as weekly for progress monitoring. The assessment correlates highly with a wide range of more time-intensive assessments and also serves as a skills diagnostic for older struggling readers.

STAR Reading—used for screening and progress-monitoring assessment—is a reliable, valid, and efficient, computer-adaptive assessment of general reading achievement and comprehension for grades 1–12. STAR Reading provides nationally norm-referenced reading scores and criterion-referenced scores. A STAR Reading assessment can be completed without teacher assistance in about 10 minutes and repeated as often as weekly for progress monitoring.

STAR Math—used for screening, progress-monitoring, and diagnostic assessment— is a reliable, valid, and efficient, computer-adaptive assessment of general math achievement for grades 1–12. STAR Math provides nationally norm-referenced math scores and criterion-referenced evaluations of skill levels. A STAR Math assessment can be completed without teacher assistance in less than 15 minutes and repeated as often as weekly for progress monitoring.

National Center on Response to Intervention
www.rti4success.org

STAR Early Literacy, **STAR Reading**, and **STAR Math** are highly rated for screening and progress monitoring by the National Center on Response to Intervention and meet all criteria for scientifically based progress-monitoring tools set by the National Center on Student Progress Monitoring.

Accelerated Math, Accelerated Reader, Advanced Technology for Data-Driven Schools, English in a Flash, MathFacts in a Flash, Renaissance, Renaissance Learning, the Renaissance Learning logo, STAR Early Literacy, STAR Math, STAR Reading, and Successful Reader are trademarks of Renaissance Learning, Inc., and its subsidiaries, registered, common law, or pending registration in the United States and other countries. AIMSweb is a trademark of Harcourt Assessment, Inc. Degrees of Reading Power (DRP) is a registered trademarks of Touchstone Applied Science Associates, Inc. DIBELS is a trademark of Dynamic Measurement Group, Inc. The Lexile Framework is a registered trademark of MetaMetrics, Inc. TPRI is a registered trademark of The Texas Education Agency. Wireless Generation® and mCLASS® are trademarks of Wireless Generation, Inc.

© 2011 by Renaissance Learning, Inc. All rights reserved. Printed in the United States of America.

This publication is protected by U.S. and international copyright laws. It is unlawful to duplicate or reproduce any copyrighted material without authorization from the copyright holder. For more information, contact:

RENAISSANCE LEARNING
P.O. Box 8036
Wisconsin Rapids, WI 54495-8036
(800) 338-4204
www.renlearn.com
answers@renlearn.com

3/11

Contents

Introduction by Dr. James R. McBride, V.P. & Chief Psychometrician, Renaissance Learning 1
The Renaissance Learning Information Pyramid 2
The Value and Cost of Information 3
Computer-Adaptive Testing 5
Item Response Theory and Its Role in CAT 5
Item Development 7
Multiple-Choice Format 7
Item Development Process 7
Items Aligned to State Standards 9
Dynamic Calibration 9
Growth Modeling 10
A Closer Look at the STAR Assessments 12
Another Way to View Data 13
About the STAR Early Literacy Assessment 14
About the STAR Reading Assessment 15
About the STAR Math Assessment 16
Reliability and Validity of the STAR Assessments 17
STAR Early Literacy Reliability and Validity 17
STAR Reading Reliability and Validity 18
STAR Math Reliability and Validity 18

Appendices

Appendix A: The STAR Assessments Relate to State and Other Assessments 19
STAR Early Literacy 19
STAR Reading 19
STAR Math 20
Correlation Summary 20
Appendix B: STAR Assessment Score Definitions 23
STAR Early Literacy 23
STAR Reading 24
STAR Math 25
Appendix C: Sample Key Reports 27
STAR Screening Report 28
STAR Student Progress Monitoring Report 30
STAR Early Literacy Student Diagnostic Report 31
STAR Annual Progress Report 32
STAR Reading Growth Report 33
STAR Math Diagnostic Report 34
STAR Math Parent Report 35
Appendix D: Sample Conversion of ATOS Levels to 100-point and 2000-Point Scales 36

Bibliography 37
References Cited 37
Key Research Support for the STAR Assessments 37
Additional Reading 39
Acknowledgements 40

Figures

Figure 1: Renaissance Learning Information Pyramid 2
Figure 2: The Value and Cost of Information 3
Figure 3: The Value of an Assessment 4
Figure 4: Illustration of a Student's Reactions to Three Test Items of Varying Difficulty 5
Figure 5: How Computer-Adaptive Technology Works 6
Figure 6: STAR Reading Goal-Setting Wizard 10
Figure 7: STAR Reports Plot a Student's Progress Toward Goals 11
Figure 8: Sample STAR Reports 12
Figure 9: STAR Learning to Read Dashboard 13
Figure 10: STAR Early Literacy Sample Items 14
Figure 11: STAR Reading Sample Items 15
Figure 12: How Comprehension Occurs 15
Figure 13: STAR Math Sample Items 16

Tables

Table 1: STAR Assessment Item Banks Administration and Breakdown, by Number and Type 8
Table B1: Summary of STAR Early Literacy Validity Studies 21
Table B2: Summary of STAR Reading Validity Studies 21
Table B3: Summary of STAR Math Validity Studies 22
Table D1: ATOS Conversion Chart to 100- and 2000-Point Scales 36

Introduction

In 1984, Judi Paul began creating quizzes for books to motvate her son, a a reluctant reader, to read more. Soon, not only were Judi's children more motivated to read but so were the neighborhood children taking Judi's quizzes. The staff at a local school heard about Judi's success, and Accelerated Reader was born.

Eight years later in 1992, Judi's husband Terry joined the company that had grown from Accelerated Reader, and he became the driving force behind the development of STAR Reading in 1996, STAR Math in 1998, and STAR Early Literacy in 2001.

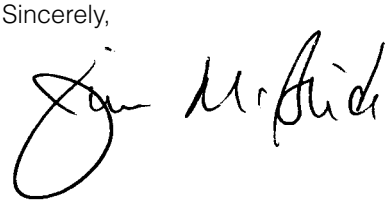
Like Accelerated Reader, the STAR assessments quickly gained favor with teachers because they are easy to use, can be administered quickly, and provide highly valid and reliable data to inform instruction at an affordable price.

Since those early kitchen-table days, Renaissance Learning has grown and evolved into the world's leading provider of computer-based assessment technology, with products in use in more than 75,000 schools, 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 Response to Intervention, the What Works Clearinghouse, and the National Center on Student Progress Monitoring.

All Renaissance Learning tools are designed to accomplish our mission—"to accelerate learning for all children and adults of all ability levels and ethnic and social backgrounds, worldwide." A key educational principle supporting this mission is the notion that "the initial step in accelerating learning is to measure its occurrence." The STAR family of assessments—STAR Early Literacy, STAR Reading, and STAR Math—does just that.

I invite you to learn more about the STAR assessments through the information enclosed. I'm confident you'll see rather quickly why teachers using Renaissance Learning tools accelerate learning, get more satisfaction from teaching, and help their students achieve higher scores on state and national tests. For additionalinformation, 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.

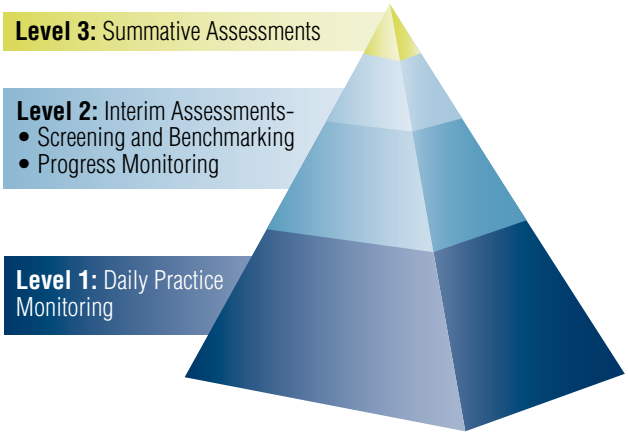
At Renaissance Learning, he has contributed to the psychometric research and development of STAR Math, STAR Reading, and STAR Early Literacy. McBride is co-editor of a leading book on the development of CAT and has authored numerous journal articles, professional papers, book chapters, and technical reports.

The Renaissance Learning Information Pyramid

All Renaissance Learning software—including the STAR assessments—runs on the web-based Renaissance Place Real Time platform, which provides a single, unified management system. Using this platform, schools and districts are able to centralize all student data from daily practice monitoring, interim (screening, benchmarking, and progress-monitoring) assessments, and summative annual state tests to create a seamless, integrated three-level assessment system. The integrated three-level assessment system was pioneered by Renaissance Learning (see Figure 1) and reflects the model experts and national educational organizations recommend (e.g., Perie, Marion, & Gong, 2007).

Renaissance Learning's *interim assessments*—the STAR Early Literacy, STAR Reading, and STAR Math assessments—make up the second, or middle, level of the Renaissance Learning Information Pyramid. The purpose of interim assessments is to determine the extent to which instruction and other daily learning tasks are strengthening students' abilities in the core academic areas and preparing them to hit end-of-year proficiency targets. These assessments are administered regularly throughout the year to help determine how all students are doing, both in groups and individually.

Figure 1: Renaissance Learning Information Pyramid



Level 2 interim assessments are generally used either for screening/benchmarking or progress monitoring. The STAR assessments, however, were developed for *both* of these purposes:[1]

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

2. **Progress-monitoring** assessments, defined as measures of academic performance by the National Center on Response to Intervention, administered more frequently than annually but less than daily—usually three to four times a year, but as often as monthly 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.

The base and topmost layers of the Renaissance Learning Information Pyramid complete an integrated, three-level system:

Level 1, *daily practice monitoring*, includes a wide variety of assessments designed to provide feedback regarding either student completion of important tasks known to improve achievement outcomes (such as reading or math problem solving) or comprehension of direct instruction—both which help to inform instruction and guide practice to improve student performance (e.g., Renaissance Learning's Accelerated Reader, Successful Reader, English in a Flash, Accelerated Math, and MathFacts in a Flash).

Level 3 *summative tests* include once-a-year, high-stakes state tests which assess student proficiency on national core and state-specific standards.

[1] Since 2009, all three STAR assessments have been highly rated by the authority on screening and progress-monitoring assessments—the National Center on Response to Intervention (NCRTI). See p. 17 for more information about the NCRTI review process.

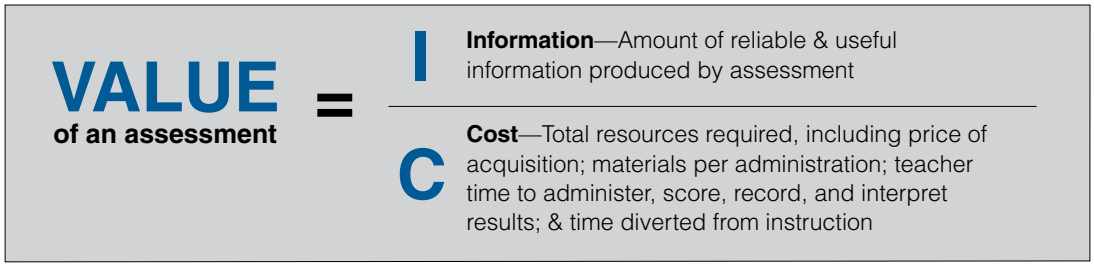
The Value and Cost of Information

When choosing an appropriate educational assessment, it is important to keep in mind the advice of the National Research Council (2008) about needs:

> Have an assessment process that is both child and teacher friendly, minimizes lost instructional time, meets the highest standards of evidence for reliability and validity for the purposes for which assessment is being planned and with the particular kinds of children that comprise the center's population, and that can be purchased and supported within the budgetary limits set out by the director. (p. 222)

Too often, schools underestimate costs by considering only the initial cash outlay for a program or system. Some solutions seem inexpensive initially but generate long-term inefficiencies and often wind up far more expensive in the long run. Two elements must be calculated: 1) the total cost of ownership, and 2) the value generated by that total cost. In the case of assessment systems, these factors constitute a return on information expressed by the formula Value = I/C (see Figure 2).

Figure 2: The Value and Cost of Information



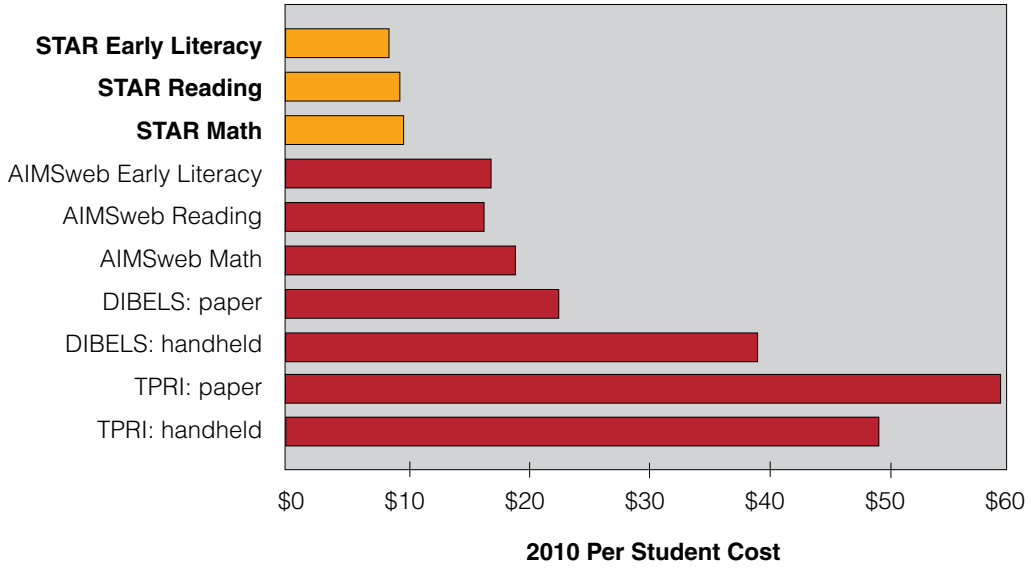
Suppose an assessment is distributed for free but requires paper administration, necessitating the duplication of test instruments, scoring sheets, record sheets, and so forth. The cost of those paper copies multiplied by the number of times that assessment will be delivered adds to the total cost of ownership. Even more significantly, if the assessment is teacher administered, the cost of that teacher's time must be figured into the calculation. A so-called one-minute probe, in reality, may occupy as many as 10 minutes, on average, of the teacher's time per student per administration (Laurits R. Christensen Associates, 2010).

The total time considered must include preparing materials, calling on the student, explaining the assessment, the administration itself, recording and entering results, and the teacher's reentry into other duties. Using the average 10-minute administration calculation, even if only three students in the classroom require testing, that may be *half an hour lost from instruction* every time the test is administered—often weekly—multiplied by the number of measures that need to be taken. As the saying goes, time is money, and teacher time is a very real cost.

This total cost, too, must be compared to the value of the information generated. If 10 minutes of testing produces only one data point on student mastery of a single skill, the return on the teacher's time is low. If the same amount of time can generate multiple data points, and/or can be applied to multiple students at the same time, the return on that same amount of time increases exponentially. A broad-based computerized assessment administered simultaneously to a whole classroom, that automatically records results in a database, provides far more information with a much higher rate of return on the teacher's time. The cost per piece of information is therefore much lower—even if the initial cost of the system is higher than the so-called free assessment.

For a practical illustration of how both parts of the Value = I/C formula work, compare curriculum-based measurements (CBM) with Renaissance Learning's STAR computer-based assessments: An independent economics research firm evaluated the annual cost of assessments frequently used for screening purposes and concluded the STAR assessments cost between one-half to one-fifth as much as the AIMSweb, DIBELS, mCLASS DIBELS, and TPRI assessments, when accounting for the value of teacher time (Laurits R. Christensen Associates, 2010). Figure 3 illustrates the comparison of average costs per student and classroom administration time for these assessments.

Figure 3: The Value of an Assessment



The figures depicted above reflect the direct costs of purchasing the assessment, the teacher's time to administer and score the individual tests if required, ongoing cost of resources if required, and the assumption of three administrations per student per year.

Source: Laurits R. Christensen Associates (2010) independent study of assessment costs.

It is also important to note that if the assessment software can be used for multiple types of assessment (e.g., both screening and diagnostic), the cost-effectiveness goes up still more. This is yet another advantage of highly efficient, multipurpose, computer-based assessments like the STAR assessments.

Computer-Adaptive Testing

STAR Early Literacy, STAR Reading, and STAR Math are all computer-adaptive tests (CATs). CATs continually adjust the difficulty of each student's test by choosing each test question based on the student's previous response. CATs save testing time and spare students the frustration of items that are too difficult and the boredom of items that are too easy.

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 good CAT uses only 25 items to yield the same information in half the time. "Adaptive tests are useful for measuring achievement because they limit the amount of time children are away from their classrooms and reduce the risk of ceiling or floor effects in the test score distribution—something that can have adverse effects on measuring achievement gains" (Agdonini & Harris, 2010, p. 215).

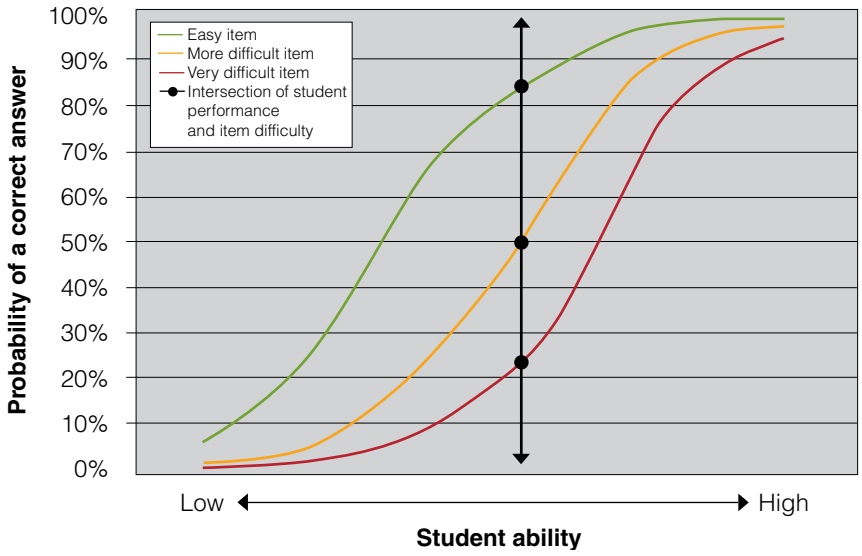
The reliability and validity of the STAR assessments has been confirmed by key federal groups including the National Center on Response to Intervention and the National Center on Student Progress Monitoring, among others (see *Reliability and Validity of the STAR Assessments*, p. 17), and is a result of the care taken by Renaissance Learning in developing each item.

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 CAT tests 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 of measuring each item's degree of difficulty and of estimating each student's achievement level from the pattern of correct and incorrect responses to items.

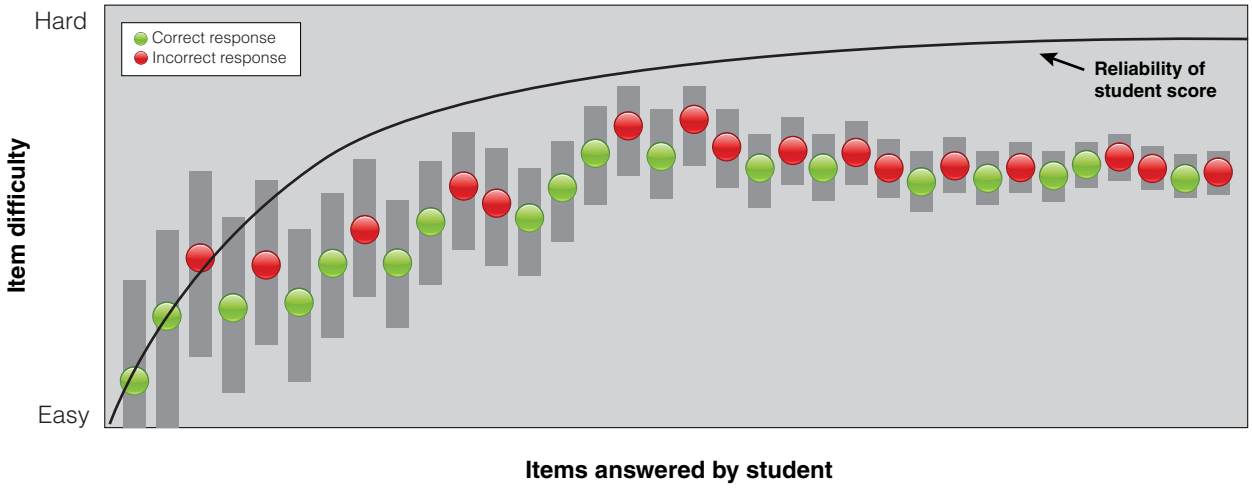
With item response theory, scientists can calculate the probability of a correct response to an item as a function of student ability. As student ability increases, so does the probability. Additionally, because some test items are harder than others, the probability trend differs from one item to another.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

Figure 4: Illustration of a Student's Reactions to Three Test Items of Varying Difficulty



In the STAR assessments, the software automatically moves up or down the scale to select questions based on the 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 assessments dynamically adjust to each student's unique responses. As a result, STAR assessments pinpoint student achievement levels quickly and efficiently. Figure 5 shows, for a single student's test, the progression of easy and more difficult items selected in a computer-adaptive assessment based on the student's previous item responses. It also shows how a computer-adaptive test's ability to select items tailored to a student helps to reduce measurement error as the test progresses.

Figure 5: How Computer-Adaptive Technology Works



Item Development

Multiple-choice format

When the STAR assessments were developed, high priority was placed on selecting a test format that was well suited to computerized testing, precise, and efficient in terms of student and teacher time. Renaissance Learning explored, researched, discussed, and prototyped several item-response formats and ultimately chose to use multiple-choice test items. Much research supports the use of the multiple-choice, 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)

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 bias, fairness, and sensitivity standards.

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."

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. Professional designers, writers, and editors, with education backgrounds and content-area expertise, develop the content for all Renaissance Learning products, including the STAR assessments. These experts follow research-based assessment item-development practices, receive ongoing item-writing and bias-and-fairness training, and adhere to the following process to ensure quality item development:

1. Analyze standards to be assessed in the categories of skill, action, vocabulary, and context; and refer to national or state resources for appropriate standard and grade-level expectation interpretation.
2. Write item specifications and provide specifications training 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 through the STAR Early Literacy, STAR Reading, and STAR Math dynamic calibration process (see *Dynamic Calibration*, p. 9).
6. Examine psychometricians' analyses of item testing results.
7. Add successful items to the operational assessment item bank.

Renaissance Learning follows strict item-writing specifications including bias and fairness criteria that 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.

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 analyses, 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 certain items for elimination, the recommendations are combined and a second review is conducted to resolve any issues.

Large item banks.
Each of the STAR assessments contains a large item bank to allow multiple administrations without risk of item overexposure. Renaissance Learning continually develops high-quality assessment items that are added to the banks to support frequent testing and achieve an even distribution of items across the difficulty levels of each STAR assessment.

The STAR assessments are fixed-length assessments, which means the item count is the sole criterion for ending a test. STAR Early Literacy and STAR Reading administer 25 items while STAR Math administers 24 items. The assessments were developed not only to provide precise measurement of student achievement in reading and math, but to do so efficiently. As mentioned earlier, computer-adaptive testing saves teachers time by automating scoring and administration. And even more importantly, it allows students to be assessed on a larger and more varied range of skills with fewer items, which results in students spending less time completing the assessment—i.e., less administration time. Table 1 shows, for each STAR test, the total number of assessment items available, number and types of items administered per testing event, and average administration time.

Table 1: STAR Assessment Item Banks and Administration Breakdown, by Number and Type

	STAR Early Literacy	STAR Reading	STAR Math
Number of Items	More than 2,400	More than 2,800	More than 1,900
Item Breakdown/ Number of Items Administered	25 items chosen from 41 skills within 7 domains: • General readiness • Graphophonemic knowledge • Phonemic awareness • Phonics • Comprehension • Structural analysis • Vocabulary	20 short comprehension items (grades 3–12) 5 extended comprehension items (grades 3–12) OR 25 short comprehension items (grades 1–2)	Items 1–8: numeration concepts Items 9–16: computation processes Items 17–24: word problems, estimation, data analysis and statistics, geometry, measurement, algebra Items cover 8 strands/ 214 objectives
Average Administration Time/Standard Deviation	8.5 minutes/ 2 minutes	7.5 minutes/ 3 minutes	11.5 minutes/ 4 minutes

Items aligned to state standards

Renaissance Learning's standards team is dedicated to best practices in standards research and alignment, based, in part, on ongoing research and consultation with leading educational organizations, including Mid-continent Research for Education and Learning (McREL), which provides research-based guidance to educators and policymakers and advocates standards-based school improvement programs; and the Northwest Regional Educational Laboratory (NWREL), which provides evaluation and research services. Renaissance standards experts, with more than 25 years of combined experience aligning assessment and practice items to state standards, analyze each assessment objective with respect to the overarching content standard and in the categories of skill, action, vocabulary, and context. They also refer to national or state resources for appropriate standards and assessment objective interpretation.

A STAR Early Literacy, STAR Reading, or STAR Math objective is aligned or developed based on whether its characteristics are the same as or a subset of the characteristics of the national or state assessment objective, which ensures assessment items do not extend beyond the domain and intent of the state assessment objective.

Dynamic Calibration

Each new STAR assessment test item goes through a calibration process to determine its exact point on the STAR difficulty scale. This is done by administering test items to large samples of students, collecting student-response and other data, and performing a statistical analysis of the response data to determine the scale values. Norming, reliability, and validity studies take place after items successfully pass through 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. In dynamic calibration, one or more new items are embedded at random points in a STAR test. These items do not count toward the student's score on the STAR assessment, but student-response data are stored for later psychometric analysis with the responses of thousands of other students. Students, on average, receive two or three additional items per test when calibration is turned on. On average, the additional calibration items increase testing time by approximately one minute.

Growth Modeling

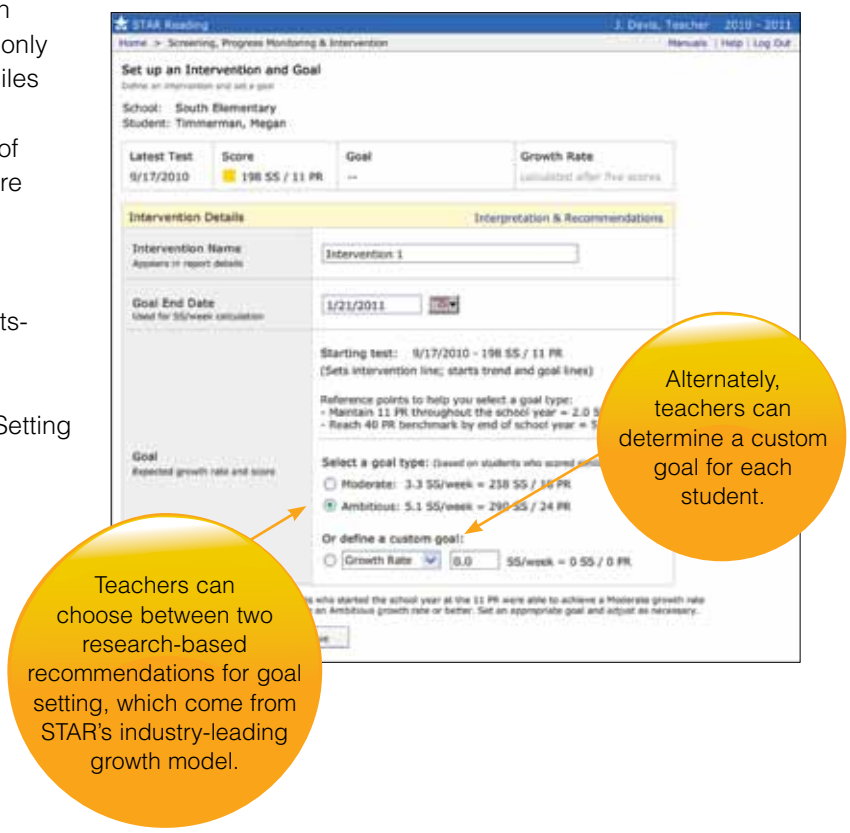
Progress monitoring is essential within a Response to Intervention framework and starts with setting appropriate goals for each student. If a progress-monitoring goal is set too high, and as a result a student does not meet that goal, the student will incorrectly appear as unable to "respond to intervention."

With STAR Early Literacy, STAR Reading, and STAR Math, educators have access to a scientific method for setting appropriate, achievable, and challenging progress-monitoring goals for students. Since thousands of schools use the STAR assessments through the web-hosted version, Renaissance Learning is able to observe how students grow. Using longitudinal data on the *learning patterns* of more than 75,000 students for early literacy, more than 1 million students for reading, and nearly 350,000 students for math,[2] the STAR assessments provide educators with critical information about how students grow over time. Specifically, the Goal-Setting Wizard in each STAR assessment uses this information to help educators set progress-monitoring goals personalized to each student—goals that are challenging but reasonable (see Figure 6).

The Renaissance Learning growth model is based on growth norms specific to each performance decile. Whereas quartiles only separate students into four groups, deciles divide students into ten groups, each representing ten percentiles. This level of specificity enables educators to compare a student's growth rate to students with scores in the same decile, making the Goal-Setting Wizard growth predictions much more accurate than a "one-size-fits-all" growth rate.

Using growth modeling data, the Goal-Setting Wizard offers research-based progress monitoring recommendations called "Moderate" and "Ambitious" goals. A moderate goal is a growth target that 50% of students nationally with the same starting score would reach. Ambitious goals are based on a rate of growth that only 25% of students in the same performance decile are able to achieve. This eliminates the need to guess how much growth constitutes good growth. With the Goal-Setting Wizard, professional judgment can now be informed by research.

Figure 6: STAR Reading Goal-Setting Wizard

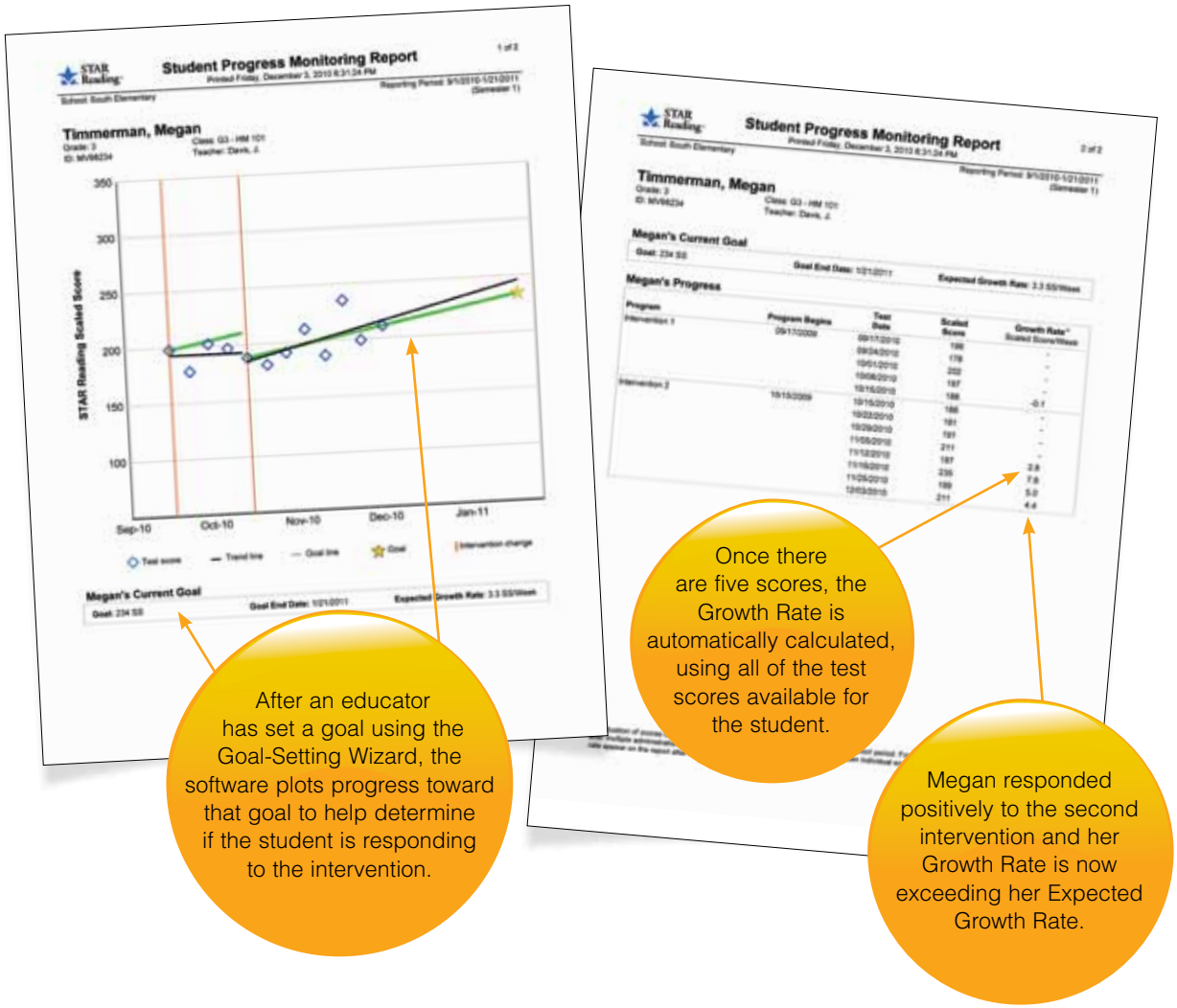


[2] All data collected by Renaissance Learning is used in complete compliance with Family Educational Rights and Privacy Act (FERPA) privacy requirements.

After a student has taken an initial STAR assessment and the teacher has selected a goal for that student, a goal line appears on the STAR Student Progress Monitoring Report. The goal line depicts the rate of growth the student must attain to meet the selected goal. Following five subsequent STAR tests, a trend line showing the student's actual growth rate is automatically drawn on the report. By comparing the goal and trend lines, educators can determine whether a student's growth trajectory is steep enough for the student to reach his goal. (See Figure 7.) Educators can then use this information to make the best instructional decisions.

The breadth and depth of our database allows us to identify the growth norms of nearly any student. Educators who use the STAR assessments have this valuable information at their fingertips, enabling them to gain a more precise understanding of how their students grow and set appropriate goals to help students reach their full potential.

Figure 7: STAR Reports Plot a Student's Progress Towards Goals



A Closer Look at the STAR Assessments

The STAR assessments allow teachers and administrators to precisely and efficiently assess student achievement in pre-reading skills (STAR Early Literacy), reading (STAR Reading), and math (STAR Math). Teachers use the wealth of data provided by the assessments to target instruction, provide students with the most appropriate instructional materials, and intervene with struggling students. Teachers access STAR assessment data via informative reports, such as those shown in Figure 8 (To view the reports in greater detail, see Appendix C, p. 27.)

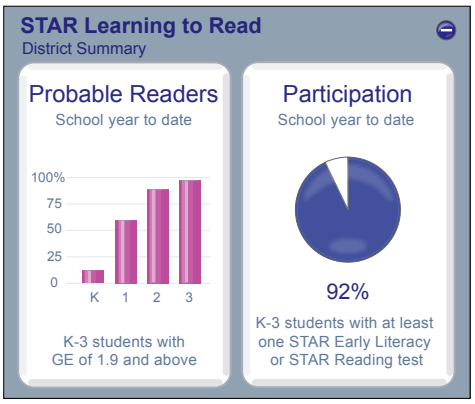
Beginning on p. 14 are brief descriptions of each assessment. For additional information, full technical manuals are available for each STAR assessment by contacting Renaissance Learning at research@renlearn.com

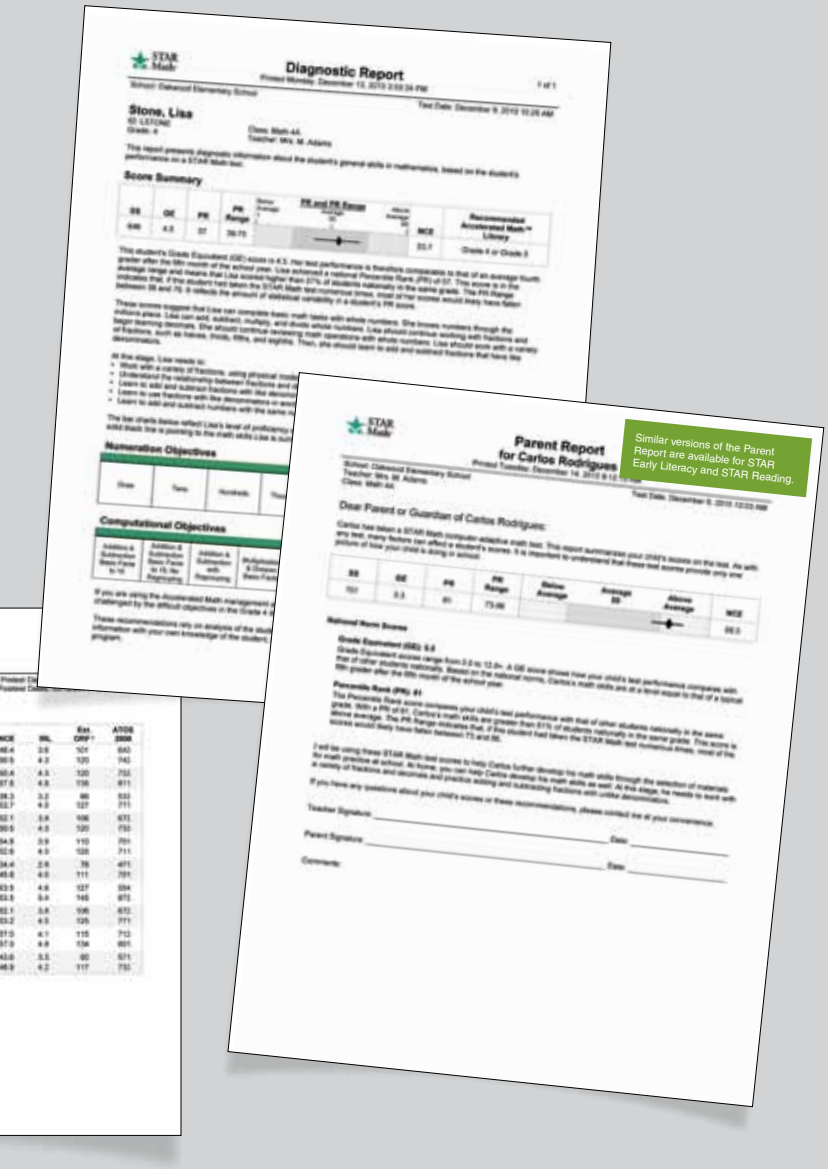
Figure 8: Sample STAR Reports

Another way to view data

In addition to the reports available in STAR Early Literacy and STAR Reading, so that teachers and administrators can zero in on their emergent readers' progress, Renaissance Learning has developed the STAR Learning to Read Dashboard (Figure 9) where teachers can view the percentage of students with STAR Early Literacy grade equivalent (GE) scores of 1.9 or above (Probable Readers) and at least one STAR Early Literacy or STAR Reading test taken school year to date (Participation).

Figure 9: STAR Learning to Read Dashboard





The **Screening Report**—complete with built-in benchmarks and cut scores—provides information about which students are responding well to core instruction and which need to be considered for intervention. (For STAR Math, it also recommends a starting point in Accelerated Math for each student.)

The **Student Progress Monitoring Report** graphs students' progress toward goals, including those set using the Goal-Setting Wizard, which was created with significant input from leading RTI experts and allows the teacher to select an appropriate goal for each student.

For STAR Early Literacy, the **Student Diagnostic Report** lists specific skills a student is struggling with and can be used for instructional planning as well as intervention decisions for struggling readers.

The **Annual Progress Report** graphs average test scores of students in a class for a school year or provides a graphic display of a student's reading or math progress across a school year in comparison to national norms. For STAR Early Literacy, the report allows comparisons of students' scores to Literacy Classifications or Risk Categories.

The **Growth Report** is used to measure progress between assessment periods.

For STAR Math, the **Diagnostic Report** shows a student's level of proficiency within numeration and computation objectives as well as a recommended starting point in Accelerated Math, which can be used for instructional planning and intervention decisions.

The **Parent Report**—available in English or Spanish—keeps the lines of communication open between school and home about student progress.

About the STAR Early Literacy assessment

The STAR Early Literacy assessment is a reliable, valid, and time-efficient assessment of early literacy skills appropriate for use within various early learning environments. Its quick and accurate results provide teachers with specific benchmarking, screening, progress-monitoring, and diagnostic information to help inform instruction to meets the needs of all students.

The development of STAR Early Literacy was based on an analysis of early learning research, with an emphasis on identifying the pre-reading and reading skills necessary for later reading success. This analysis revealed seven major content areas (Adams, 1990; Anderson, Hiebert, Scott, & Wilkinson, 1985; Anderson, Wilson, & Fielding, 1988; National Reading Panel, 2000; Snow, Burns, & Griffin, 1998; Trelease, 1995) that became the basis for the seven skill domains assessed in STAR Early Literacy: general readiness, graphophonemic knowledge, phonemic awareness, phonics, comprehension, structural analysis, and vocabulary.

The STAR Early Literacy domains include four of the five critical areas of reading instruction identified by the National Reading Panel. While the fifth area identified—fluency—is not directly assessed in STAR Early Literacy, it is highly correlated with other reading skills such as comprehension. Because fluency is an important component of general reading achievement, STAR Early Literacy provides an Estimated Oral Reading Fluency score for beginning readers (grades 1–3).

Figure 10: STAR Early Literacy Sample Items

"Listen carefully. The pictures are king, fish, foot. Click on the picture that has a different beginning sound than the others."

"Read the story. Then pick the word that means about the same as dirty."

Estimated oral reading fluency is an estimation of the number of words a student should be able to read correctly on a grade-level appropriate passage within a one-minute time span. The score is based on research linking STAR Early Literacy and STAR Reading scores to student performance on the DIBELS oral reading fluency measure. Students with high oral reading fluency demonstrate accurate decoding, automatic word recognition, and appropriate use of the rhythmic aspects of language (e.g., intonation, phrasing, pitch, emphasis).

Renaissance Learning also examined the early learning research to determine both the skills to assess within the seven selected domains and the design of the emergent reader test items. In total, 41 skills sets (containing a total of 147 skills) were identified. The test items were designed to incorporate text, graphics, and audio, as appropriate, to assess the skills in the most effective way possible, and the instructions were written to be explicit, clear, and consistent from item to item so that students would be able to test independently. Figure 10 shows sample assessment items.

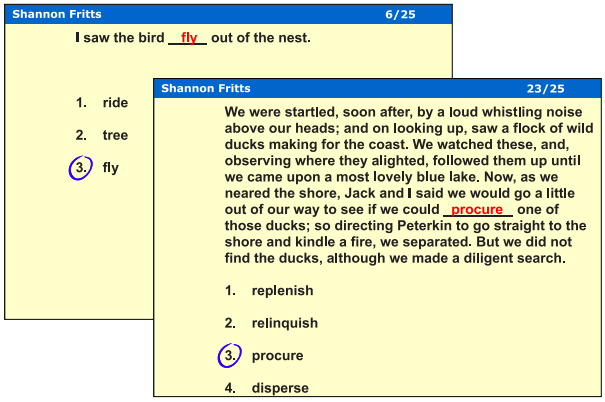
Using STAR Early Literacy data.

STAR Early Literacy is used for screening/benchmarking and progress monitoring of emergent readers in grades pre-K–3. The assessment also provides diagnostic data to make instructional decisions and help identify likely gaps in knowledge for students experiencing reading difficulties. See Appendix B, p. 23, for an explanation of all scores reported by STAR Early Literacy.

About the STAR Reading assessment

The STAR Reading assessment is a reliable, valid, and time-efficient assessment of general reading comprehension appropriate for use with a variety of instructional and curriculum frameworks. Its quick and accurate results provide teachers with specific benchmarking, screening, and progress-monitoring information to help tailor instruction, monitor reading growth, and improve reading achievement for all students.

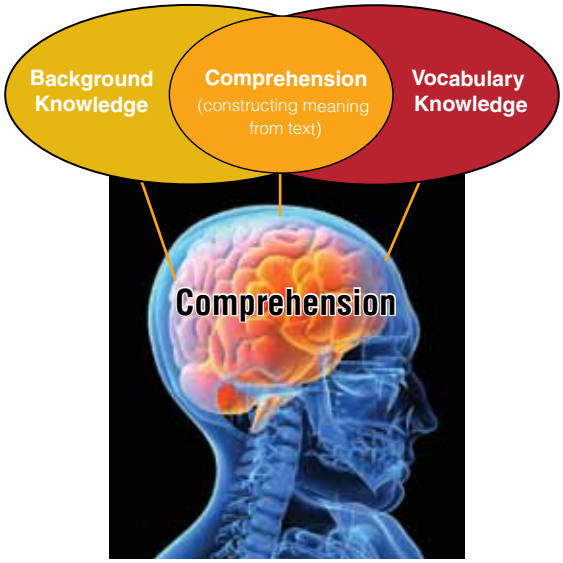
Figure 11: STAR Reading Sample Items



STAR Reading assesses reading comprehension through the use of two item types: short comprehension items and extended comprehension items (see Figure 11). The use of the former is based on abundant and long-standing research verifying that vocabulary is closely tied to comprehension (Davis, 1942; Just & Carpenter, 1987; see Figure 12). STAR Reading's short comprehension items contain one complete contextual sentence with a tightly controlled vocabulary level and a single-word deletion. The longer extended comprehension items contain multiple sentences and a single-word deletion.

For both, the information needed to determine the correct answer is given within the assessment-item passage provided, with the semantics and syntax of each context sentence arranged to provide clues to the correct answer choice. The only prior knowledge needed is an understanding of the words in the text passages and answer choices. The items require reading comprehension because the student must actually interpret the meaning of the sentence or passage to choose the correct answer; all answer choices "fit" the context sentence either semantically or syntactically but only one is correct. The reading levels of the items range from kindergarten through post high school.

Figure 12: How Comprehension Occurs



STAR Reading results for students in grades 1–4 include an Estimated Oral Reading Fluency score (for more information, see p. 14). Although fluency is not directly assessed in STAR Reading, it is highly correlated with reading comprehension and an important component of general reading achievement.

Using STAR Reading data.

STAR Reading is used for screening/benchmarking and progress monitoring of students in grades 1–12.[3] It automates benchmarks, cut scores, progress-monitoring goals, and instructional recommendations, and helps the teacher determine if student achievement is heading in the right direction. One score reported by STAR Reading is a student's ZPD, or individualized reading range. To experience optimal growth, the student chooses books with ATOS readability levels within this range.[4] See Appendix B, p. 24, for an explanation of all scores reported by STAR Reading.

[3] STAR Reading may also be used with kindergarten students, though the assessment has not been normed for this age group. Students are ready to test in STAR Reading when they have a 100-word sight vocabulary, which for some students may not occur until later in first grade.

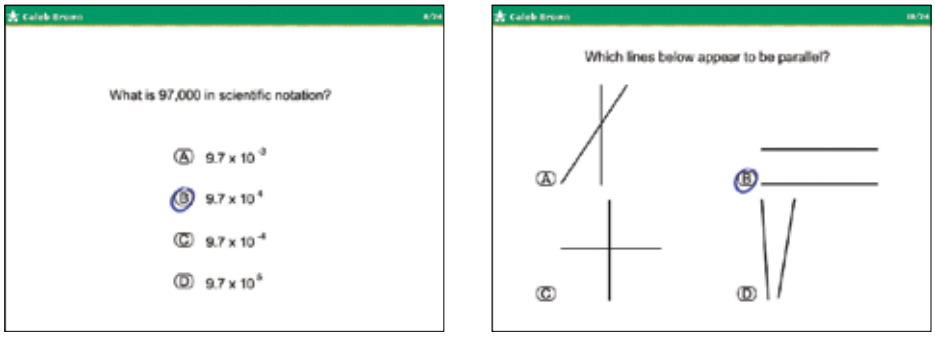
[4] To help educators, Renaissance Learning has developed conversion charts that correlate ATOS readability levels with a 100-point scale like the one used by Degrees of Reading Power and a 2000-point scale like the one used by The Lexile Framework. See Appendix D, p. 36, for a sample of this conversion. The full conversion table is available in *The Development of ATOS: The Renaissance Learning Readability Formula* (Milone, 2009): http://doc.renlearn.com/KMNet/R004250827GJ11C4.pdf

About the STAR Math assessment

The STAR Math assessment is a reliable, valid, and time-efficient assessment of mathematics skills appropriate for use within various instructional and curriculum frameworks. Its quick and accurate results provide teachers with specific benchmarking, screening, progress-monitoring, and diagnostic information to help tailor instruction, monitor math growth, and improve math achievement for all students.

The content for STAR Math is based on analysis of state standards, various curriculum materials, test frameworks, and content-area research, including best practices for mathematics instruction. Research indicates that numeration concepts are key for deep conceptual development and that computational processes emphasizing fluency complement conceptual development. STAR Math provides a unique system of joint analysis of numeration and computational processes in addition to content for geometry, measurement, algebra, data analysis and statistics, estimation, and word problems. The STAR Math item bank includes 214 core math objectives, with multiple items available to measure each objective. Figure 13 shows sample assessment items.

Figure 13: STAR Math Sample Items



Using STAR Math data.

STAR Math is used for screening/benchmarking, progress monitoring, and diagnosis of students' skills in grades 1–12.[5] See Appendix B, p. 25, for an explanation of all scores reported by STAR Math.

[5] STAR Math may also be used with kindergarten students, though the assessment has not been normed for this age group. Students are ready to test in STAR Math when they have a 100-word sight vocabulary, which for some students may not occur until later in first grade.

Reliability and Validity of the STAR Assessments

In 2009, the U.S. Department of Education began funding the National Center on Response to Intervention (NCRTI),[6] 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, the STAR assessments have maintained the NCRTI's highest ratings, meaning the assessments fulfill both key elements of a school's RTI framework.

RTI stands for Response to Intervention, a framework for making instructional decisions based on data, in order to accelerate learning for all students. Interim assessments, like STAR Early Literacy, STAR Reading, and STAR Math, play a key role in RTI, helping to provide data to inform and improve instruction.

National Center on Response to Intervention
www.rti4success.org

STAR Early Literacy, STAR Reading, and STAR Math are highly rated for screening and progress monitoring by the National Center on Response to Intervention.

The NCRTI's positive review of the STAR assessments confirms the reliability and validity of each test, and is in agreement with other assessment experts (Salvia, Ysseldyke, & Bolt, 2010). Reliability is the extent to which a test yields consistent results from one test administration to another. To be reliable, tests must yield consistent results. The validity of an assessment is the degree to which it measures what it is intended to measure and is often used to judge a test's effectiveness. Standard error of measurement (SEM) measures the precision of a test score. It provides a means to gauge the extent to which scores would be expected to fluctuate because of imperfect reliability, which is a characteristic of all educational tests.

The following provides a brief explanation of the reliability and validity of each STAR assessment. For more information, see *Appendix A: The STAR Assessments Relate to State and Other Assessments*, p. 19, for tables detailing the reliability and validity studies conducted for each assessment.

STAR Early Literacy reliability and validity

STAR Early Literacy's reliability was estimated using three different methods (split-half, generic, and test-retest) to determine the overall precision of its test scores. The analysis was based on test results from more than 9,000 students. The reliability estimates were very high, comparing favorably with reliability estimates typical of other published early literacy tests.

For STAR Early Literacy to measure 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 validity research study of STAR Early Literacy in spring 2001 to assess reliability, validity, and score distributions by age and grade. Although the validity research study sample was targeted to include schools using certain standardized early literacy and reading assessments, the participating school districts, specific schools, and individual students were approximately representative of the U.S. school population in terms of the following three key variables: 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 validity of STAR Early Literacy as a tool for assessing early literacy skills. For more information, see *Appendix A: The STAR Assessments Relate to State and Other Assessments*, p. 19.

[6] Prior to 2009, the U.S. Department of Education's National Center on Student Progress Monitoring (NCSPM) reviewed the STAR assessments and determined each fit all criteria for scientifically based progress-monitoring tools (http://www.studentprogress.org/chart/docs/print_chart122007.pdf).

STAR Reading reliability and validity

STAR Reading's reliability was estimated using three different methods (split-half, generic, and test-retest) when the test was first normed in spring 1999 with a sample of 30,000 students from 269 schools in 47 U.S. states. Schools and districts were selected based on their geographic location, per-grade district enrollment, and socioeconomic status. The reliability estimates were very high, comparing favorably with reliability estimates typical of other published reading tests.

For STAR Reading to measure reading achievement, Renaissance Learning knew it was necessary that its scores correlate highly with other measures of reading achievement. To that end, during the STAR Reading norming study, schools submitted their students' STAR Reading results along with data on how their students performed on a wide variety of other popular standardized tests. Scores were received for more than 10,000 students. The resulting correlation estimates were substantial and reflect well on the validity of STAR Reading as a tool for assessing reading achievement. (For more information, see *Appendix A: The STAR Assessments Relate to State and Other Assessments*, p. 19.) Additional data supporting the validity of STAR Reading are collected and reported on a continuing basis, resulting in a large and growing body of validity evidence that now includes hundreds of validity studies.

In spring 2008, STAR Reading was renormed, using national samples of students drawn from routine administrations of STAR Reading. In other words, the students in the 2008 norming sample took STAR Reading tests as they are administered in everyday use. This was a change from the previous special-purpose norming study, in which national samples of schools were cast, and those schools were administered a special norming version of the assessment. In total, 69,738 students in grades 1–12 were part of the 2008 norming study, representing 2,709 schools across 48 U.S. states and the District of Columbia.

STAR Math reliability and validity

STAR Math reliability was estimated using three different methods (split-half, generic, and test-retest) when the test was normed in the spring of 2002. Renaissance Learning obtained a nationally representative sample by selecting school districts and schools based on their geographic location, per-grade district enrollment, and socioeconomic status. The final norming sample for STAR Math included approximately 29,200 students from 312 schools in 48 U.S. states. The reliability estimates were very high, comparing favorably with reliability estimates typical of other published math achievement tests.

For STAR Math to measure math achievement, Renaissance Learning knew it was necessary that its scores correlate highly with other measures of math achievement. During the STAR Math norming study, schools submitted their students' STAR Math results along with data on how their students performed on other popular standardized tests. Scores were received for more than 10,000 students. The resulting correlation estimates were substantial and reflect well on the validity of STAR Math as a tool for assessing math achievement. (For more information, see *Appendix A: The STAR Assessments Relate to State and Other Assessments*, p. 19.) As with STAR Reading, additional data supporting the validity of STAR Math are collected and reported on a continuing basis, resulting in a large and growing body of validity evidence that now includes hundreds of validity studies.

Appendix A: The STAR Assessments Relate to State and Other Assessments

STAR Early Literacy

Studies have been conducted with STAR Early Literacy and the following state assessments to correlate or statistically link* the tests:

- Florida Comprehensive Assessment Test (FCAT)
- Indiana Statewide Testing for Educational Progress (ISTEP)
- Iowa Test of Basic Skills (ITBS)

STAR Early Literacy scores also correlate with several assessments of early literacy skills:

- AIMSweb
- 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)
- Gates-MacGinitie Reading Test (GMRT)
- Group Reading Assessment and Diagnostic Evaluation (GRADE)
- Metropolitan Readiness Test (MRT)
- Michigan Literacy Progress Profile (MLPP)
- NWEA Levels Test
- Running Records
- Stanford Achievement Test (SAT-9)
- Texas Primary Reading Inventory (TPRI)

STAR Reading

Studies have been conducted with STAR Reading and the following state assessments to correlate or statistically link* the tests:

- Alabama Reading and Mathematics Test (ARMT)
- Arizona's Instrument to Measure Standards (AIMS)
- Arkansas Augmented Benchmark Examinations (AABE)
- California Achievement Test (CAT)
- California Standards Tests for English-Language Arts and Mathematics (CST)
- Colorado Student Assessment Program (CSAP)
- Delaware Student Testing Program (DSTP)—Reading
- Florida Comprehensive Assessment Test (FCAT)
- Georgia's Criterion-Referenced Competency Tests (CRCT)
- Idaho Standards Achievement Tests (ISAT)
- Illinois Standards Achievement Test (ISAT)–Reading
- Indiana Statewide Testing for Education Progress-Plus Assessments (ISTEP+)
- Iowa Test of Basic Skills (ITBS)
- Kansas Reading Assessment
- Kansas State Assessment Program (KSAP)
- Kentucky Core Content Test (KCCT)
- Louisiana Educational Assessment Program (LEAP); Integrated Educational Assessment Program Assessments (iLEAP)
- Michigan Educational Assessment Program (MEAP)—English Language Arts

* 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 the STAR assessments, email research@renlearn.com

- Michigan Educational Assessment Program (MEAP)—Reading
- Mississippi Curriculum Test, Second Edition (MCT2)
- Missouri Assessment Program (MAP) Grade-Level Assessments
- New Jersey Assessment of Skills and Knowledge (NJ ASK)
- New Mexico Standards Based Assessments (SBA) for Reading and Math
- New York State Assessment Program (NYSTP) for English Language Arts
- North Carolina End-of-Grade Tests (NC EOG)
- Ohio Achievement Assessments (OAA)
- Oklahoma Core Curriculum Tests (OCCT)
- South Carolina's Palmetto Assessment of State Standards (PASS) for English Language Arts and Mathematics
- Tennessee Comprehensive Assessment Program (TCAP) Achievement Test
- Texas Assessment of Knowledge and Skills (TAKS)
- Utah's Criterion-Referenced Test for English Language Arts
- Washington's Measurements of Student Progress (MSP)

STAR Math

Studies have been conducted with STAR Math and the following state assessments to correlate or statistically link* the tests:

- Alabama Reading and Mathematics Test (ARMT)
- Arizona's Instrument to Measure Standards (AIMS)
- Arkansas Augmented Benchmark Examinations (AABE)
- California Achievement Test (CAT)
- California Standards Tests for English-Language Arts and Mathematics (CST)
- Colorado Student Assessment Program (CSAP)
- Connecticut Mastery Test (CMT)
- Delaware Student Testing Program (DSTP)—Mathematics
- Florida Comprehensive Assessment Test (FCAT)
- Georgia's Criterion-Referenced Competency Tests (CRCT)
- Georgia High School Graduation Tests (GHSGT)
- Idaho Standards Achievement Tests (ISAT)
- Indiana Statewide Testing for Education Progress-Plus Assessments (ISTEP+)
- Iowa Test of Basic Skills (ITBS)
- Kansas Mathematics Assessment
- Kansas State Assessment Program (KSAP)
- Kentucky Core Content Test (KCCT)
- Louisiana Educational Assessment Program (LEAP); Integrated Educational Assessment Program Assessments (iLEAP)
- Michigan Educational Assessment Program (MEAP)—Mathematics
- Minnesota Comprehensive Assessments
- Mississippi Curriculum Test, Second Edition (MCT2)
- Missouri Assessment Program (MAP) Grade-Level Assessments
- New Mexico Standards Based Assessments (SBA) for Reading and Math
- North Carolina End-of-Grade Test (NC EOG)
- Ohio Achievement Assessments (OAA)
- Oklahoma Core Curriculum Tests (OCCT)
- Pennsylvania System of School Assessment (PSSA)

* 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 the STAR assessments, email research@renlearn.com

- Rhode Island New Standards Reference Exams (NSRES)—Mathematics
- South Carolina's Palmetto Assessment of State Standards (PASS) for English Language Arts and Mathematics
- Tennessee Comprehensive Assessment Program (TCAP) Achievement Test
- Texas Assessment of Knowledge and Skills (TAKS)
- Virginia Standards of Learning (SOL) Tests
- Washington Assessment of Student Learning (WASL)
- Washington's Measurements of Student Progress (MSP)

Correlation summary

Tables B1, B2, and B3 summarize the reliability and validity studies conducted for each STAR assessment. For additional information, full technical manuals are available for each STAR assessment by contacting Renaissance Learning at research@renlearn.com

Table B1: Summary of STAR Early Literacy Validity Studies

	Predictive			Concurrent		
Grade	**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

Table B2: Summary of STAR Reading Validity Studies

	Predictive			Concurrent and Other External Validity		
Grade	**Studies**	**Students**	**Average Correlation**	**Studies**	**Students**	**Average Correlation**
1	6	74,877	.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

Table B3: Summary of STAR Math Validity Studies

	Predictive			Concurrent		
Grade	**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

Appendix B: STAR Assessment Score Definitions

STAR Early Literacy

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 performance and oral reading fluency.

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. 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 to 100, is criterion-referenced and represents the percentage of items a student would be expected to answer correctly within the following seven domains, covering 41 literacy skills:

General readiness (GR): Ability to identify shapes, numbers, colors, and patterns; explore word length and word pairs; and examine oral and print numbers.

Graphophonemic knowledge (GK): Ability to relate letters to corresponding sounds; addresses skills like matching upper- and lowercase letters, recognizing the alphabet, naming letters, recognizing letter sounds, and knowing alphabetical order.

Phonemic awareness (PA): Ability to detect and identify individual sounds within spoken words. Assesses skills like rhyming words; blending word parts and phonemes; discriminating between beginning, medial, and ending sounds; understanding word length; and identifying missing sounds.

Phonics (PH): Ability to read words by using the sounds of letters, letter groups, and syllables. Addresses skills like identifying short and long vowels, beginning and ending consonants, and consonant blends and digraphs; recognizing word families; and using strategies such as consonant and vowel replacement.

Comprehension (CO): Ability to understand what has been read aloud, understand word meaning, and read text correctly. Addresses skills like identifying and understanding words, selecting the word that best completes a sentence, and answering items about stories.

Structural analysis (SA): Ability to understand the structure of words and word parts. Addresses skills like finding words, adding beginning or ending letters or syllables to a word, building words, and identifying compound words.

Vocabulary (VO): Ability to identify high-frequency words, match pictures with synonyms, match words with phrases, match stories with words, identify opposites, match pictures with opposite word meanings, and identify opposite word meanings.

Scaled score (SS) is useful in comparing student performance over time and is calculated based on the difficulty of items and the number of correct responses. Because the same range is used for all students, scaled scores are also useful for comparing student performance across grade levels. STAR Early Literacy scaled scores range from 300 to 900 and relate directly to the literacy classifications above.

Skill set score, ranging from 0 to 100, is criterion-referenced and estimates a student's percent of mastery of specific skills within the seven domains listed above.

STAR Reading

Grade equivalent (GE) score, ranging 0.0 to 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.

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 performance and oral reading fluency.

Grade placement (GP) is a numeric representation of a student's grade level, based on the month in which a student takes a STAR Reading assessment. The STAR Reading software considers the standard school year from September through June and assigns increment values of 0.0 through 0.9 to these months. The software automatically assigns a student's grade placement using grade level and the month in which that student takes a STAR Reading assessment. GP is important because percentile rank and normal curve equivalent values are based on a student's scaled score *and* the student's grade placement at the time of the test.

Instructional reading level (IRL) is a criterion-referenced score that indicates the highest reading level at which a student is at least 80 percent proficient at recognizing words and understanding material with instructional assistance. For example, a seventh-grade student with an IRL score of 8.0 reads eighth-grade words with 80 percent accuracy or better. IRL scores are Pre-Primer (PP), Primer (P), grades 1.0 through 12.9, and Post-High School (PHS).

Normal curve equivalent (NCE) score, ranging from 1 to 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. NCEs are mostly used for research and are useful in making comparisons between different achievement tests and for statistical computations—for example, determining an average score for a group of students.

Percentile rank (PR) score, ranging from 1 to 99, is norm-referenced and provides the best measure of a student's level of reading achievement 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.

Percentile rank range (PR Range) is norm-referenced and reflects the amount of statistical variability in a student's percentile rank score. For example, a student with a percentile rank range of 32–59 is likely to score within that range if the STAR Reading assessment is taken again within a short time frame—for example, 4 to 6 weeks.

Scaled score (SS) is useful in comparing student performance over time and is calculated based on the difficulty of itemsand the number of correct responses. Because the same range is used for all students, scaled scores are also useful for comparing student performance across grade levels. STAR Reading scaled scores range from 0 to 1400. All norm-referenced scores are derived from the scaled score.

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

STAR Math

Accelerated Math Library Recommendation is determined based on the results of a student's STAR Math assessment and helps educators place a student in the Accelerated Math library that will be of the most benefit, based on that student's individual achievement level.

Grade equivalent (GE) score, ranging 0.0 to 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.

Grade placement (GP) is a numeric representation of a student's grade level, based on the month in which a student takes a STAR Math assessment. The STAR Math software considers the standard school year from September through June and assigns increment values of 0.0 through 0.9 to these months. The software automatically assigns a student's grade placement using grade level and the month in which that student takes a STAR Math assessment. GP is important because percentile rank and normal curve equivalent values are based on a student's scaled score *and* the student's grade placement at the time of the test.

Math instructional level (MIL) is a student's current level of math instruction. Teachers can adjust this value to enable the software to raise or lower the beginning difficulty level of the first STAR Math assessment a student takes.

Normal curve equivalent (NCE) score, ranging from 1 to 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. NCEs are mostly used for research and are useful in making comparisons between different achievement tests and for statistical computations—for example, determining an average score for a group of students.

Percentile rank (PR) score, ranging from 1 to 99, is norm-referenced and provides the best measure of a student's level of math achievement 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.

Percentile rank range (PR Range) is norm-referenced and reflects the amount of statistical variability in a student's percentile rank score. For example, a student with a percentile rank range of 32–59 is likely to score within that range if the STAR Math assessment is taken again within a short time frame—for example, 4 to 6 weeks.

Scaled score (SS) is useful in comparing student performance over time and is calculated based on the difficulty of items and the number of correct responses. Because the same range is used for all students, scaled scores are also useful for comparing student performance across grade levels. STAR Math scaled scores range from 0 to 1400. All norm-referenced scores are derived from the scaled score.

Appendix C: Sample Key Reports

On the following pages are full-size samples of the reports shown on pp. 12 and 13.[7]

The complete menu of reports available for each STAR assessment is found in a separate publication from Renaissance Learning, *Key Report Samples*, available online from http://doc.renlearn.com/KMNet/R003563228GE7E80.pdf or by request to (800) 338-4204.

[7] Reports are regularly reviewed and may vary from those shown as enhancements are made.

The Screening Report is available for STAR Early Literacy, STAR Reading, and STAR Math.

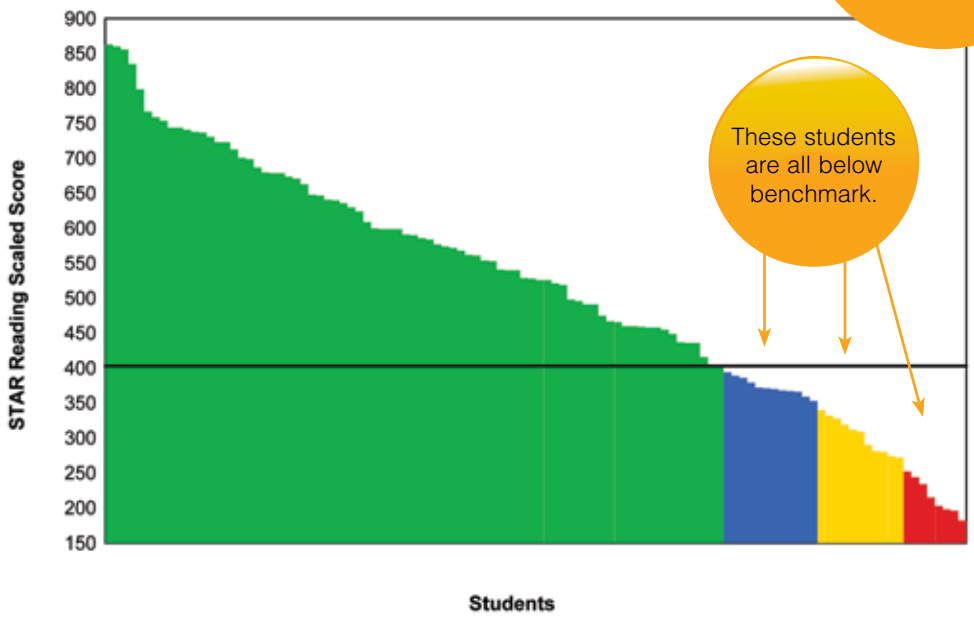
Screening Report

Printed Tuesday, September 7, 2010 3:21:18 PM

y School Reporting Peri

Page 1 of the screening report shows a graphical representation of all students in the grade.

Grade: 4



Categories / Levels	Benchmark Scores		Students	
	Percentile Rank	Scaled Score	Number	Percent
At/Above Benchmark	At/Above 40 PR	At/Above 403	79	72%
– Categories Below Benchmark				
On Watch	Below 40 PR	Below 403	12	11%
Intervention	Below 25 PR	Below 347	11	10%
Urgent Intervention	Below 10 PR	Below 267	8	7%
Students Tested			110	

Key questions to ask based on this and other information:
Are you satisfied with the number of students at or above benchmark (green)? Which students represented by blue are you "worried about" and what support within or beyond core instruction is warranted? What support is needed for students represented by yellow? Do all students represented by red need urgent intervention?

Use these key questions to help determine next steps.

RTI

Screening is the first step in Response to Intervention (RTI). Use this report for grade-level planning and identifying students who need the most help.

STAR Early Literacy

Screening Report

Printed Friday, September 17, 2010 3:56:07 PM

2 of 7

School: West Elementary School

Pages 2 and beyond provide a breakdown of which students are in each category.

Grade: 1

Urgent Intervention

Student	Class	Teacher	Test Date	SS	Est. ORF#	Literacy Skills Classification
Mulder, Barry	G1 (Paulson)	Paulson, L.	09/17/2010	548	6	Late Emergent Reader
Grady, Lee	G1 (Davos)	Davos, M.	09/16/2010	546	6	Late Emergent Reader
Janowitsch, Pamela	G1 (Hansen)	Hansen, M.	09/14/2010	544	6	Late Emergent Reader
Peterson, Jason C.	G1 (Davos)	Davos, M.	09/17/2010	540	5	Late Emergent Reader
Hong, Luoluo	G1 (Davos)	Davos, M.	09/17/2010	536	4	Late Emergent Reader

STAR Reading

Screening Report

Printed Tuesday, September 7, 2010 3:21:18 PM

2 of 5

School: Oakwood Elementary School

Provides a suggested independent reading range for each student to optimize reading growth.

Grade: 4

Urgent Intervention

Student	Class	Teacher	Test Date	SS	GP	PR	GE	Est. ORF#	ZPD
Brennan, Kenneth	Grade 4 (Dillon)	Dillon, A.	09/07/2010	182	4.0	2	1.9	45	1.9-2.9
Franklin, Kurt	Grade 4 (Dillon)	Dillon, A.	09/07/2010	196	4.0	3	2.0	47	2.0-3.0
Cochran, Herman	Grade 4 (Dillon)	Dillon, A.	09/07/2010	198	4.0	3	2.0	47	2.0-3.0
Pace, Bennie	Grade 4 (Dillon)	Dillon, A.	09/07/2010	203	4.0	3	2.0	48	2.0-3.0
Bradshaw, Delores	Grade 4 (Dillon)	Dillon, A.	09/07/2010	215	4.0	4	2.1	50	2.1-3.1

STAR Math

Screening Report

Printed Friday, September 11, 2009 4:50:07 PM

2 of 7

School: East Elementary

Period: 9/08/2009-9/11/2009 (Fall Screening)

STAR Math recommends an appropriate instructional level for each student.

Grade: 5

Urgent Intervention

Student	Class	Teacher	Test Date	SS		PR	GE	Recommended Accelerated Math™ Library
Helgestad, Tammy	G-5 Marshall	Travis, C.	09/10/2009	499	5.03	4	2.6	Grade 2 or Grade 3
Hughes, Jenny	G-5 Davis	Davis, J.	09/09/2009	501	5.03	5	2.6	Grade 2 or Grade 3
Behl, Brenda	G-5 Travis	Travis, C.	09/09/2009	507	5.03	5	2.7	Grade 2 or Grade 3
Friske, Cheryl	G-5 Reynolds	Reynolds, P.	09/10/2009	507	5.03	5	2.7	Grade 2 or Grade 3
Thompson, Charmaine	G-5 Jefferson	Jefferson, D.	09/09/2009	514	5.03	6	2.8	Grade 3
Fauerbach, Diane	G-5 Reynolds	Reynolds, P.	09/11/2009	515	5.03	6	2.8	Grade 3

The Student Progress Monitoring Report is available for STAR Early Literacy, STAR Reading, and STAR Math.

RTI

Page 1 of this report graphs a student's scores in relation to their goal, giving the teacher a picture of the student's progress.

Student Progress Monitoring Report

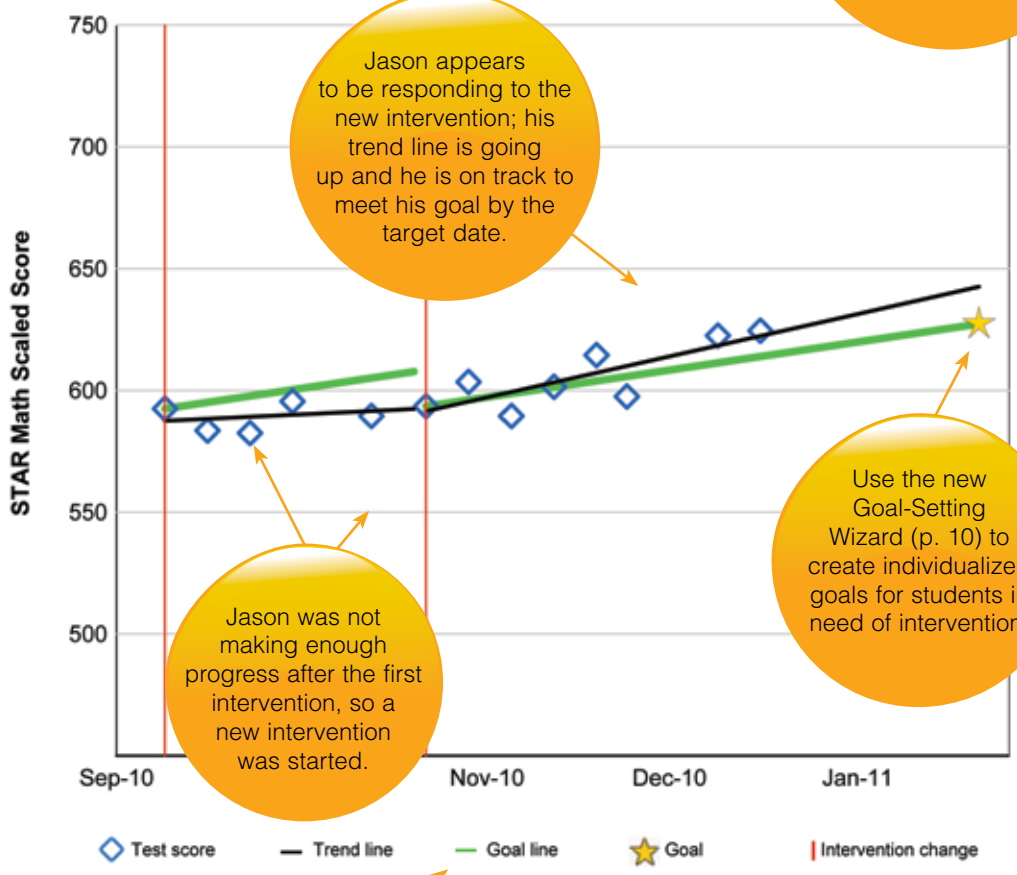
Printed Friday, December 17, 2010 4:53:06 PM

Reporting P

Peterson, Jason

Grade: 5
ID: JV234243897

Class: G5 - Travis
Teacher: Travis, C.



Jason appears to be responding to the new intervention; his trend line is going up and he is on track to meet his goal by the target date.

Jason was not making enough progress after the first intervention, so a new intervention was started.

Use the new Goal-Setting Wizard (p. 10) to create individualized goals for students in need of intervention.

Jason's Current Goal

Goal: 627 SS	**Goal End Date:** 1/21/2011	**Expected Growth Rate:** 2.6 SS/Week

While the goal line projects an intervention outcome, the trend line shows the student's actual progress toward that goal.

STAR Early Literacy™

Student Diagnostic Report
Skill Set Scores

Printed Tuesday, June 7, 2011 9:17:05 AM

School: Oakwood Elementary Scho

Reporting Period: 09/01/
(2010-20

Use this report as a part of a plan to personalize instruction for Robert.

Report Options
Group By: Class
Sort By: ID

Identifies whether the student is an emerging reader, transitional reader, or probable reader.

Estada, Robert

Class: Mrs. Rowley's Class
Teacher: Mrs. C. Rowley
Grade: 1
Test Date: 06/06/2011
Score: 685 SS (Scaled Score)

Literacy Classification			
Early Emergent Reader	Late Emergent Reader	**Transitional Reader**	Probable Reader
SS 300-487	SS 488-674	**SS 675-774**	SS 775-900

Domain	Score
General Readiness	85
Graphophonemic Knowledge	74
Phonemic Awareness	64
Comprehension	61
Phonics	59
Vocabulary	58
Structural Analysis	53

Skill Sets Within Each Literacy Domain

Skill set scores, ranging from 0-100, estimate the student's percent of mastery of skills in each set.

General Readiness	Skill Set Score
Differentiating letters	93
Differentiating words from letters	92
Comparing word length (written)	91
Differentiating shapes	90
Recognizing position words	86
Matching numbers and objects	82
Differentiating word pairs	82
→ Completing sequences	69
→ Identifying word boundaries	67

Graphophonemic Knowledge	
Naming letters	91
Matching upper- and lowercase letters	85
Recognizing letter sounds	85
→ Recognizing alphabetic sequence	66
→ Using alphabetical order	60

Phonemic Awareness	
Blending word parts	83
→ Blending phonemes	74
→ Comparing word length (oral)	64
→ Discriminating beginning and ending sounds	62
→ Identifying rhyming words	59
→ Identifying missing sounds	47

Comprehension	
→ Reading and understanding words	70
→ Reading and completing sentences	59
→ Reading and understanding paragraphs	47

Phonics	Skill Set Score
Replacing beginning and ending consonants	79
→ Matching sounds within word families	69
→ Identifying consonant blends	68
→ Substituting consonant sounds	63
→ Identifying beginning consonant sounds	
→ Identifying consonant digraphs	
→ Replacing vowels	
→ Identifying medial short vowels	
→ Identifying ending consonant sounds	
→ Identifying medial long vowels	
Matching and recognizing short vowel sounds	
Matching and recognizing long vowel sounds	34

Robert needs additional practice in these skills.

Vocabulary	
Matching words and pictures	96
→ Recognizing antonyms	56
→ Recognizing synonyms	53

Structural Analysis	
→ Finding words	58
→ Building words	55
→ Identifying compound words	49

Estimated Oral Reading Fluency
Words Correct Per Minute: 24 Est. ORF is available for tests taken in grades 1-3.

→ Next Steps: These are skill sets the student is ready to learn and practice, based on their Scaled Score. Skill sets with a score below 40 may not have been presented to the student yet or may be too difficult at this time.

The Annual Progress Report is available in STAR Early Literacy, STAR Reading, and STAR Math.

Annual Progress Report

Printed Tuesday, June 7, 2011 10:55:01AM

ary School

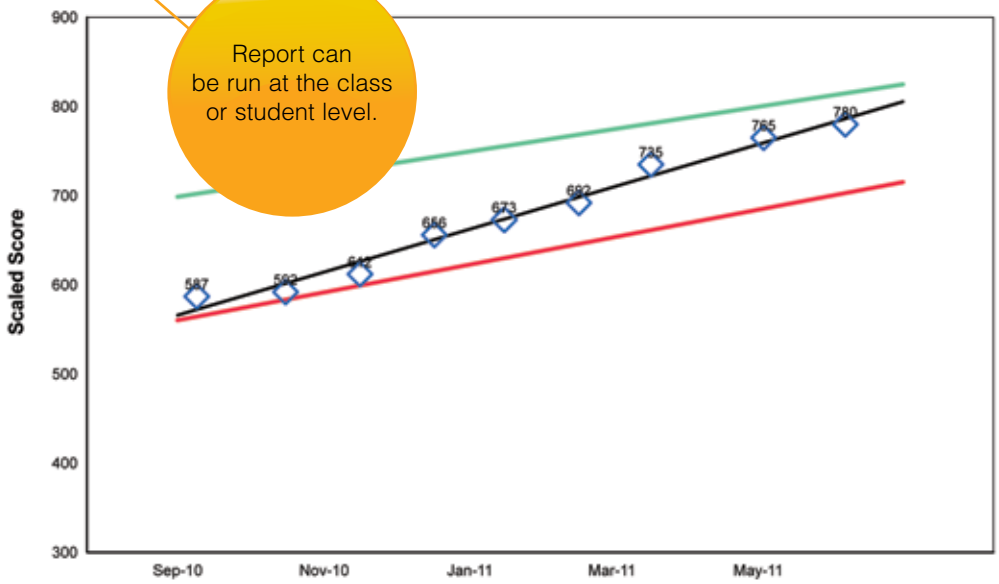
Reporting Period: 9/1/2010 - 6/10/2011

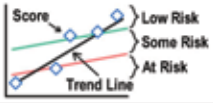
Report Options
Reporting Parameter Group: All Demographics [Default]
Group by: Class
Comparison: Risk Categories

Class: Mrs. Rowley's Class

Grade: 1
Teacher: Rowley, C.

Report can be run at the class or student level.





Diamonds (blue) show scores for each STAR assessment in the school year. For two or more scores, a trend line (black) is displayed. The graph is separated into At Risk, Some Risk, and Low Risk for acquiring reading skills. For additional information, see *STAR Early Literacy Technical Manual*, found in the software.

Test	Date Range	Number of Students	GP	Scaled Sc	Literacy Domain Scores						
					GR	GK	PA	CO	PH	VO	SA
1	09/01/2010 - 09/30/2010	15	1.04			60	47	45	43	43	35
2	10/01/2010 - 10/31/2010	15	1			0	47	45	43	43	36
3	11/01/2010 - 11/30/2010	15					51	48	47	46	39
4	12/01/2010 - 12/31/2010	15					58	56	54	53	47
5	01/01/2011 - 01/30/2011	15					62	59	57	56	50
6	02/01/2011 - 02/28/2011	15					65	62	60	59	54
7	03/01/2011 - 03/31/2011	15					73	70	68	66	63
8	04/01/2011 - 04/30/2011	15				4	84	75	73	72	69
9	06/01/2011 - 05/31/2011	15	1.8			87	81	79	77	75	73

This table provides additional detail about each of the testing ranges during the school year.

GP: Grade Placement
GR: General Readiness
GK: Graphophonemic Knowledge
PA: Phonemic Awareness
CO: Comprehension
PH: Phonics
VO: Vocabulary
SA: Structural Analysis

STAR Reading

Growth Report

Printed Tuesday, May 17, 2011 10:15:21 AM

School: Oakwood Elementary Sc

2010 - 09/10/2010
/2011 - 05/13/2011

Use this report to measure students' progress between two testing periods.

Estimated Oral Reading Fluency is an estimation of the number of words a student should be able to read within a one-minute time span.

A similar version of the Growth Report is available for STAR Math.

Class: Grade 4 (Adams)

		Teacher	Test Date	GP	SS	GE	PR	PR Range	NCE	IRL	Est. ORF[b]	ATOS 2000
	4 (Adams)	Adams, Marcie	09/09/2010	4.02	436	3.8	47	41-49	48.4	3.6	101	643
			05/13/2011	4.84	522	4.8	51	46-58	50.5	4.3	120	743
	4 (Adams)	Adams, Marcie	09/10/2010	4.03	521	4.8	69	64-72	60.4	4.3	120	733
			05/13/2011	4.84	586	5.5	64	61-69	57.5	4.8	135	811
Bosley, Matthew	Grade 4 (Adams)	Adams, Marcie	09/10/2010	4.03	370	3.1	29	25-32	38.3	3.2	86	533
			05/13/2011	4.84	550	5.1	57	45-61	53.7	4.5	127	771
Chang, Michelle	Grade 4 (Adams)	Adams, Marcie	09/09/2010	4.02	460	4.2	54	45-59	52.1	3.8	106	672
			05/13/2011	4.84	520	4.8	51	43-56	50.5	4.3	120	733
Gonzales, Maria	Grade 4 (Adams)	Adams, Marcie	09/10/2010	4.03	475	4.4	59	49-62	54.8	3.9	110	701
			05/13/2011	4.84	542	5.0	55	49-60	52.6	4.5	125	711
Halden, Susan	Grade 4 (Adams)	Adams, Marcie	09/09/2010	4.02	340	2.7	23	20-26	34.4	2.9	78	471
			05/13/2011	4.84	481	4.4	42	37-46	45.8	4.0	111	701
O'Neil, Sarah	Grade 4 (Adams)	Adams, Marcie	09/10/2010	4.03	554	5.2	74	68-79	63.5	4.6	127	554
			05/13/2011	4.84	646	6.1	74	72-80	63.5	5.4	145	872
Richmond, Angela	Grade 4 (Adams)	Adams, Marcie	09/10/2010	4.03	458	4.1	54	48-58	52.1	3.8	106	672
			05/12/2011	4.84	545	5.1	56	52-59	53.2	4.5	125	771
Rodrigues, Carlos	Grade 4 (Adams)	Adams, Marcie	09/10/2010	4.03	496	4.6	63	55-65	57.0	4.1	115	712
			05/12/2011	4.84	581	5.5	63	59-67	57.0	4.8	134	801
Stone, Lisa	Grade 4 (Adams)	Adams, Marcie	09/08/2010	4.02	394	3.4	38	30-39	43.6	3.3	92	571
			05/12/2011	4.84	508	4.7	48	44-50	48.9	4.2	117	733

Summary

	GP	SS	GE	PR	NCE	IRL	Est. ORF[b]	ATOS 2000
Pretest Mean	4.03	452	4.0	52	51.1			630
Posttest Mean	4.84	548	5.1	56	53.2			771
Change	+0.81	+96	+1.1	+4	+2.1			+141
Number of Students: 10								

Susan has shown excellent growth this year; she's catching up to her peers.

ATOS 2000 is a correlation between Renaissance Learning ATOS levels and a 2000-point scale similar to The Lexile Framework.

[b]Est. ORF: Estimated Oral Reading Fluency is only reported for tests taken in grades 1-4.

RTI

This report provides information about Lisa's math skills based on her STAR Math test results.

STAR Math

Diagnostic Report

Printed Monday, December 13, 2010 3:55:34 PM

School: Oakwood Elementary School

Test Date: Dec

Stone, Lisa

ID: LSTONE
Grade: 4

Class: Math 4A
Teacher: Mrs. M. Adams

This report presents diagnostic information about the student's general skills in mathematics, based on the student's performance on a STAR Math test.

Score Summary

SS	GE	PR	PR Range	PR and PR Range (Below Average 1 / Average 50 / Above Average 99)	NCE	Recommended Accelerated Math™ Library
646	4.5	57	38-75		53.7	Grade 4 or Grade 5

This student's Grade Equivalent (GE) score is 4.5. Her test performance is therefore comparable to that of an average fourth grader after the fifth month of the school year. Lisa achieved a national Percentile Rank (PR) of 57. This score is in the average range and means that Lisa scored higher than 57% of students nationally in the same grade. The PR Range indicates that, if this student had taken the STAR Math test numerous times, most of her scores would likely have fallen between 38 and 75. It reflects the amount of statistical variability in a student's PR score.

These scores suggest that Lisa can complete basic math tasks with whole numbers. She knows numbers through the millions place. Lisa can add, subtract, multiply, and divide whole numbers. Lisa should continue working with fractions and begin learning decimals. She should continue reviewing math operations with whole numbers. Lisa should work with a variety of fractions, such as halves, thirds, fifths, and eighths. Then, she should learn to add and subtract fractions that have like denominators.

At this stage, Lisa needs to:

- Work with a variety of fractions, using physical models to enhance understanding
- Understand the relationship between fractions and decimals
- Learn to add and subtract fractions with like denominators
- Learn to use fractions with like denominators in word problems
- Learn to add and subtract numbers with the same number of decimal places

These are the next concepts Lisa should learn.

The bar charts below reflect Lisa's level of proficiency within the Numeration and Computation ou . The solid black line is pointing to the math skills Lisa is currently developing.

Numeration Objectives

Ones	Tens	Hundreds	Thousands	Hundred Thousands	Fractions & Decimals	Advanced Concepts I	Advanced Concepts II

Computational Objectives

Addition & Subtraction Basic Facts to 10	Addition & Subtraction Basic Facts to 18, No Regrouping	Addition & Subtraction with Regrouping	Multiplication & Division Basic Facts	Advanced Computation with Whole Numbers	Fractions & Decimals I	Fractions & Decimals II	Percents, Ratios & Proportions	Multiplication & Division of Mixed Numbers

If you are using the Accelerated Math management software system with Lisa, assign the Grade 4 library. If she is not challenged by the difficult objectives in the Grade 4 library, move her to the Grade 5 library.

These recommendations rely on analysis of the student's performance on one STAR Math test. Please combine this ation with your own knowledge of the student, and use your professional judgment when designing an instructional

Lisa's Accelerated Math library recommendation.

Similar versions of the Parent Report are available for STAR Early Literacy and STAR Reading.

STAR Math

Parent Report
for Carlos Rodrigues

Printed Tuesday, December 14, 2010 9:12:15 AM

School: Oakwood Elementary School
Teacher: Mrs. M. Adams
Class: Math 4A

Test Date: December 9, 2010 10:03 AM

Dear Parent or Guardian of Carlos Rodrigues:

Carlos has taken a STAR Math computer-adaptive math test. This report summarizes your child's scores on the test. As with any test, many factors can affect a student's scores. It is important to understand that these test scores provide only one picture of how your child is doing in school.

SS	GE	PR	PR Range	Below Average	Average 50	Above Average	NCE
701	5.5	81	73-86				68.5

National Norm Scores

Grade Equivalent (GE): 5.5
Grade Equivalent scores range from 0.0 to 12.9+. A GE score shows how your child's test performance compares with that of other students nationally. Based on the national norms, Carlos's math skills are at a level equal to that of a typical fifth grader after the fifth month of the school year.

Percentile Rank (PR): 81
The Percentile Rank score compares your child's test performance with that of other students nationally in the same grade. With a PR of 81, Carlos's math skills are greater than 81% of students nationally in the same grade. This score is above average. The PR Range indicates that, if this student had taken the STAR Math test numerous times, most of his scores would likely have fallen between 73 and 86.

I will be using these STAR Math test scores to help Carlos further develop his math skills through the selection of materials for math practice at school. At home, you can help Carlos develop his math skills as well. At this stage, he needs to work with a variety of fractions and decimals and practice adding and subtracting fractions with unlike denominators.

If you have any questions about your child's scores or these recommendations, please contact me at your convenience.

Teacher Signature: _____ Date: _____

Parent Signature: _____ Date: _____

Comments:

This report is available in English or Spanish.

This report helps keep parents involved by giving recommendations on what skills the student needs to work on.

Appendix D: Sample Conversion of ATOS Levels to 100-Point and 2000-Point Scales

Table D1 shows a sample of the conversion chart developed by Renaissance Learning to correlate ATOS levels with a 100-point scale like the one used by Degrees of Reading Power and a 2000-point scale like the one used by The Lexile Framework.

The full conversion table is available in *The Development of ATOS: The Renaissance Learning Readability Formula* (Milone, 2009): http://doc.renlearn.com/KMNet/R004250827GJ11C4.pdf

Table D1: ATOS Conversion Chart to 100- and 2000-Point Scales

ATOS Levels	ATOS 100-Point Scale Values	ATOS 2000-Point Scale Values
0.3	21	23
0.5	27	31
0.7	30	39
1.0	34	51
1.3	37	120
1.5	38	157
1.7	40	190
2.0	41	241
2.3	43	301
2.5	44	348
2.7	44	381
3.0	46	440
3.3	47	491
3.5	47	530
3.7	48	561
4.0	49	619
4.3	50	669
4.5	50	701
4.7	51	731
5.0	52	781
5.3	53	829
5.5	53	860
5.7	54	888
6.0	55	921
6.3	56	951
6.5	57	979

ATOS Levels	ATOS 100-Point Scale Values	ATOS 2000-Point Scale Values
6.7	57	999
7.0	58	1029
7.3	60	1061
7.5	60	1080
7.7	61	1098
8.0	61	1121
8.3	62	1149
8.5	62	1161
8.7	62	1178
9.0	63	1201
9.3	63	1237
9.5	64	1258
9.7	64	1272
10.0	65	1293
10.3	65	1314
10.5	65	1328
10.7	66	1342
11.0	66	1364
11.3	66	1385
11.5	67	1399
11.7	67	1413
12.0	67	1434
12.3	68	1455
12.5	68	1469
12.7	68	1483
13.0	69	1504

Bibliography

References cited

Adams, M. J. (1990). *Beginning to read*. London: MIT Press.

Agdonini, R., & Harris, B. (2010). An experimental evaluation of four elementary school math curricula. *Journal of Research on Educational Effectiveness, 3*(3), 199–253.

Anderson, R. C., Hiebert, E. H., Scott, J. A., & Wilkinson, I. A. G. (1985). *Becoming a nation of readers: The report on the commission of reading*. Washington, DC: The National Institute of Education.

Anderson, R. C., Wilson, P. T., & Fielding, L. G. (1988). Growth in reading and how children spend their time outside of school. *Reading Research Quarterly, 23*, 285–303.

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.

Davis, F. B. (1942). Two new measures of reading ability. *Journal of Educational Psychology, 33*, 365–372.

Hafner, L. E. (1966). Cloze procedure. *Journal of Reading, 9*(6), 415–421.

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.

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

Just, M. A., & Carpenter, P. A. (1987). *The psychology of reading and language comprehension*. Boston: Allyn & Bacon.

Laurits R. Christensen Associates. (2010). *A cost analysis of early literacy, reading, and mathematics assessments: STAR, AIMSweb, DIBELS, and TPRI.* Madison, WI: Author. Available online from http://doc.renlearn.com/KMNet/R003711606GF4A4B.pdf

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

Mattimore, P. (2009, February 5). *Why our children need national multiple choice tests*. Retrieved August 18, 2009, 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.

Milone, M. (2009). *The development of ATOS: The Renaissance readability formula.* Wisconsin Rapids, WI: Renaissance Learning, Inc. Available online from http://doc.renlearn.com/KMNet/R004250827GJ11C4.pdf

National Reading Panel (2000). *Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction*. Bethesda, MD: Author.

National Research Council. (2008). *Early childhood assessment: Why, what, and how.* Committee on Developmental Outcomes and Assessments for Young Children, C. E. Snow & S. B. Van Hemel (Eds.). Board on Children, Youth, and Families, Board on Testing and Assessment, Division of Behavioral and Social Sciences and Education. Washington, DC: The National Academies Press.

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.

Perie, M., Marion, S., & Gong, B. (2007). *A framework for considering interim assessments*. Dover, NH: National Center for the Improvement of Educational Assessment. Retrieved June 21, 2007, from http://www.nciea.org/publications/ConsideringInterimAssess_MAP07.pdf

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.

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.

Salvia, J., Ysseldyke, J., & Bolt, S. (2010). *Assessment: In special and inclusive education* (11th ed.). Belmont, CA: Wadsworth Publishing.

Snow, C. E., Burns, M. E., & Griffin, P. (1998). *Preventing reading difficulties in young children*. Washington, DC: National Academy Press.

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

Trelease, J. (1995). *The read-aloud handbook*. New York: Penguin Books.

Key research support for the STAR assessments

STAR Early Literacy.

Borman, G., & Dowling, M. (2004). *Assessment committee analysis of reading assessment measures coding form: STAR Early Literacy (SEL).* Madison: University of Wisconsin, Department of Educational Leadership and Policy Analysis. Available online from http://www.education.wisc.edu/elpa/people/faculty/Borman/AssessmentCommitteeCodingFormSEL.pdf

Renaissance Learning. (2009). *STAR Early Literacy: Technical manual*. Wisconsin Rapids, WI: Author. Available from Renaissance Learning by request to research@renlearn.com

Salvia, J., Ysseldyke, J., & Bolt, S. (2007). STAR Early Literacy computer-adaptive diagnostic assessment. In *Assessment: In special and inclusive education* (10th ed., pp. 439–440). Boston: Houghton Mifflin.

U.S. Department of Education: National Center on Response to Intervention. (2010). *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. (2009). *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 Student Progress Monitoring. (2006). *Review of progress monitoring tools* [Review of STAR Early Literacy]. Washington, DC: Author. Available online from http://www.studentprogress.org/chart/docs/print_chart122007.pdf

STAR Reading.

Renaissance Learning. (2009). *STAR Reading: Technical manual*. Wisconsin Rapids, WI: Author. Available from Renaissance Learning by request to research@renlearn.com

Salvia, J., Ysseldyke, J., & Bolt, S. (2010). Using technology-enhanced assessments: STAR Reading. In *Assessment: In special and inclusive education* (11th ed., pp. 330–331). Belmont, CA: Wadsworth Publishing.

U.S. Department of Education: National Center on Response to Intervention. (2010). *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. (2009). *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. (2006). *Review of progress monitoring tools* [Review of STAR Reading]. Washington, DC: Author. Available online from http://www.studentprogress.org/chart/docs/print_chart122007.pdf

STAR Math.

Renaissance Learning. (2009). *STAR Math: Technical manual*. Wisconsin Rapids, WI: Author. Available from Renaissance Learning by request to research@renlearn.com

Salvia, J., Ysseldyke, J., & Bolt, S. (2010). Using technology-enhanced assessments: STAR Math. In *Assessment: In special and inclusive education* (11th ed., pp. 329–330). Belmont, CA: Wadsworth Publishing.

U.S. Department of Education: National Center on Response to Intervention. (2010). *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. (2009). *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 Student Progress Monitoring. (2006). *Review of progress monitoring tools* [Review of STAR Math]. Washington, DC: Author. Available online from http://www.studentprogress.org/chart/docs/print_chart122007.pdf

Additional reading

Betts, J. (2007, May). *Developmental trajectories of early literacy skills predict later reading problems*. Poster presented at the annual meeting of the Association for Psychological Science, Washington, DC.

Betts, J., Good, R., Cummings, K., Williams, K., Hintze, J., & Ysseldyke, J. (2007, February). *Psychometric adequacy of measures of early literacy skills*. Symposium on the assessment of early literacy presented at the meeting of the National Association of School Psychologists, New York, NY.

Betts, J., & McBride, J. (2008, February). *Investigating construct validity of four measures of early literacy skills*. Paper presented at the meeting of the National Association of School Psychologist, New Orleans, LA.

Betts, J., & McBride, J. (2008, March). *Using computerized adaptive testing and an accelerated longitudinal design to index learning progressions in early mathematics development*. Paper presented at the meeting of the American Education Research Association, New York, NY.

Betts, J., & McBride, J. (2008, July). *Investigating the measurement equivalence and construct validity of tests of early reading skills.* Poster presented at the meeting of the Society for the Scientific Study of Reading, Asheville, NC.

Betts, J., & McBride, J. (2009, February). *From conceptual to concrete: Creating coherent & balanced assessment systems: Predictive power of interim assessments*. Paper presented at the winter meeting of the Council of Chief State School Officers (CCSSO), State Collaborative on Assessment and Student Standards (SCASS), Technical Issues in Large-Scale Assessments (TILSA), Orlando, FL.

Betts, J., Topping, K., & McBride, J. (2007, April). *An international linking study of a computerized adaptive test of reading with a traditional paper-and-pencil test of reading comprehension*. Paper presented at the meeting of the National Council on Measurement in Education, Chicago, IL.

McBride, J., & Betts, J. (2007, June). *Eleven years of assessing K–12 achievement using CAT: STAR Reading, STAR Math, and STAR Early Literacy.* Paper presented at GMAC Computerized Adaptive Testing Conference, Minneapolis, MN.

McBride, J., Ysseldyke, J., Milone, M., & Stickney, E. (2010). Technical adequacy and cost benefit of four measures of early literacy. *Canadian Journal of School Psychology, 25*(2), 189–204.

Ysseldyke, J., Burns, M. K., Scholin, S. E., & Parker, D. C. (2010). Instructionally valid assessment within Response to Intervention. *Teaching Exceptional Children, 42*(4), 54–61.

Ysseldyke, J., & McLeod, S. (2007). Using technology tools to monitor Response to Intervention. In S. R. Jimerson, M. K. Burns, & A. M. VanDerHeyden (Eds.), *Handbook of Response to Intervention: The science and practice of assessment and intervention* (pp. 396–407). New York: Springer.

Acknowledgements

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

Thomas P. Hogan, Ph.D., is a professor of psychology and a Distinguished University Fellow at the University of Scranton. He has more than 40 years of experience conducting reviews of mathematics curricular content, principally in connection with the preparation of a wide variety of educational tests, including the Stanford Diagnostic Mathematics Test, Stanford Modern Mathematics Test, and the Metropolitan Achievement Test. Hogan has published articles in the *Journal for Research in Mathematics Education* and *Mathematical Thinking and Learning*, and he has authored two textbooks and more than 100 scholarly publications in the areas of measurement and evaluation. He has also served as consultant to a wide variety of school systems, states, and other organizations on matters of educational assessment, program evaluation, and research design.

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.

Michael Milone, Ph.D., is a research psychologist and 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.

Sharif M. Shakrani, Ph.D., is 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 eight 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).

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. Professor Ysseldyke's 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. Ysseldyke 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

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

Thomas P. Hogan, Ph.D., is a professor of psychology and a Distinguished University Fellow at the University of Scranton.

Sharif M. Shakrani, Ph.D., is 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.

James R. McBride, Ph.D., is vice president and chief psychometrician for Renaissance Learning.

James Ysseldyke, Ph.D., is Emma Birkmaier Professor of Educational Leadership in the Department of Educational Psychology at the University of Minnesota.

Michael Milone, Ph.D., is a research psychologist and award-winning educational writer and consultant to publishers and school districts.

Renaissance Learning™
Advanced Technology for Data-Driven Schools™

L1526.0611.FP.7.5M
R14807