

THE RESEARCH FOUNDATION FOR
STAR Assessments™

The Science of STAR™



STAR™
Assessments

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

All logos, designs, and brand names for Renaissance Learning's products and services, including but not limited to Accelerated Maths, Accelerated Reader, Renaissance Learning, STAR, STAR Assessments, STAR Early Literacy, STAR Maths and STAR Reading are trademarks of Renaissance Learning, Inc., and its subsidiaries, registered, common law, or pending registration in the United States and other countries. All other product and company names should be considered the property of their respective companies and organizations.

© 2015 by Renaissance Learning, Inc. All rights reserved. Printed in the United Kingdom.

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 UK LTD
32 Harbour Exchange Square,
London E14 9GE

T: 020 7184 4000
www.renlearn.co.uk
answers@renlearn.co.uk

National Center on
INTENSIVE INTERVENTION

at American Institutes for Research ■

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



National Center on Response to Intervention
www.rti4success.org

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

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

Contents

Quick-reference guide to the STAR Assessments	5.
Letter to Educators from Jim McBride, Vice President and Chief Psychometrician	6.
STAR overview	7.
Learning Progressions	8.
Framework for Learning Progressions in reading and maths	10.
The Renaissance Learning Information Pyramid	12.
The Value and Cost of Information	13.
Computer Adaptive Testing	14.
Growth Modelling	18.
STAR Record Book	21.
A closer look at STAR Assessments	22.
Reliability and Validity of the STAR Assessments	25.
Instructional planning	27.
STAR Assessment score definitions	28.
Conclusions	31.
Bibliography	32.

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 critical early literacy domains. A STAR Early Literacy assessment can be completed without teacher assistance in about 10 minutes by emergent readers 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 across all school years. STAR Reading provides nationally norm-referenced reading scores and criterion-referenced scores. A STAR Reading assessment can be completed without teacher assistance in about 15 minutes and repeated as often as weekly for progress monitoring.

STAR Maths - used for screening, progress-monitoring, and diagnostic assessment - is a reliable, valid, and efficient, computer-adaptive assessment of general maths achievement across all school years. STAR Maths provides nationally norm-referenced maths scores and criterion-referenced evaluations of skill levels. A STAR Maths assessment can be completed without teacher assistance in less than 15 minutes and repeated as often as weekly for progress monitoring.

Introduction

Dear Educator,

Renaissance Learning™ is the world's leading provider of computer-based assessment technology, with products in use worldwide covering all primary and secondary Years. Renaissance Learning tools have a research base unmatched by makers of other educational products and have met the highest review standards set by reputable organizations such as the National Center on Intensive Intervention, the National Center on Response to Intervention, National Center on Student Progress Monitoring, the National Dropout Prevention Center, the Promising Practices Network, and the What Works Clearinghouse and the National Foundation for Educational Research (NFER).

All Renaissance Learning tools are designed to accomplish our mission - "accelerating learning for all." A key educational principle supporting this mission is the notion that "the initial step in accelerating learning is to measure its occurrence." Our assessments - STAR Early Literacy, STAR Reading, and STAR Maths - do just that.

There is a reason over 30,000 schools worldwide use at least one STAR™ assessment. They quickly gain favor with educators because of their ease of use, quick administration times, and ability to provide teachers with highly valid and reliable data upon completion of each test. The computer-based STAR assessment system is a multipurpose tool. STAR is used for screening and progress monitoring, and also includes resources that target instruction for all kinds of learners. Students who are most at risk can be identified quickly. No time is wasted in diagnosing their needs, allowing intervention to begin immediately.

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

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



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 Maths, 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.

STAR overview

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

Three STAR assessments measure student achievement in four areas:

- STAR Early Literacy assesses early literacy and early numeracy skills
- STAR Reading assesses reading skills
- STAR Maths assesses maths skills

All STAR assessments include skills-based test items, Learning Progressions, and in-depth reports. Operating on the Renaissance Place hosted platform, STAR assessments are a comprehensive assessment system for data-driven schools. The assessments provide accurate data in a short amount of time by combining computer-adaptive technology with a specialised psychometric test design that utilised item response theory (IRT).

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

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

Table 1: Research support for STAR assessments

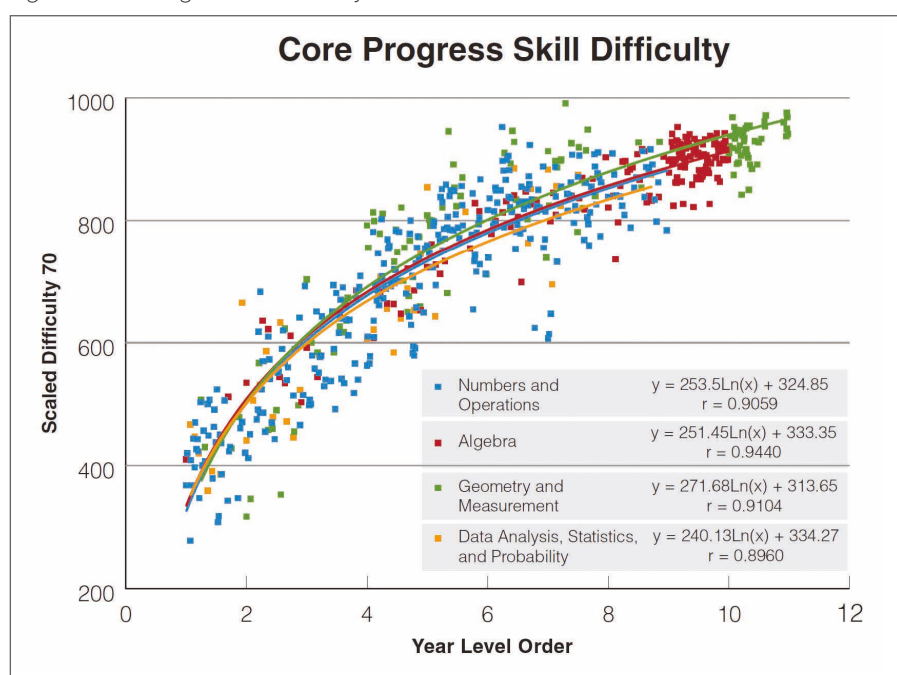
Assessment	Total research publications	Independent research publications
STAR Early Literacy	21	14
STAR Reading	76	22
STAR Maths	65	21

Learning Progressions

STAR is built upon Learning Progressions. A Learning Progression takes years to develop through a continuous process of research, expert review, and iterative revision. Continually refined since 2007, learning progressions are an interconnected web of prerequisite skills.

Once built, the Core Progress skills were field tested through the STAR assessments. The results were remarkable. As illustrated in the graph below, the order of skills in Core Progress are highly correlated with the difficulty level of STAR assessment items. With a strong correlation, the natural next step was to statistically link Core Progress to STAR assessments.

Figure 1: Core Progress Skill Difficulty

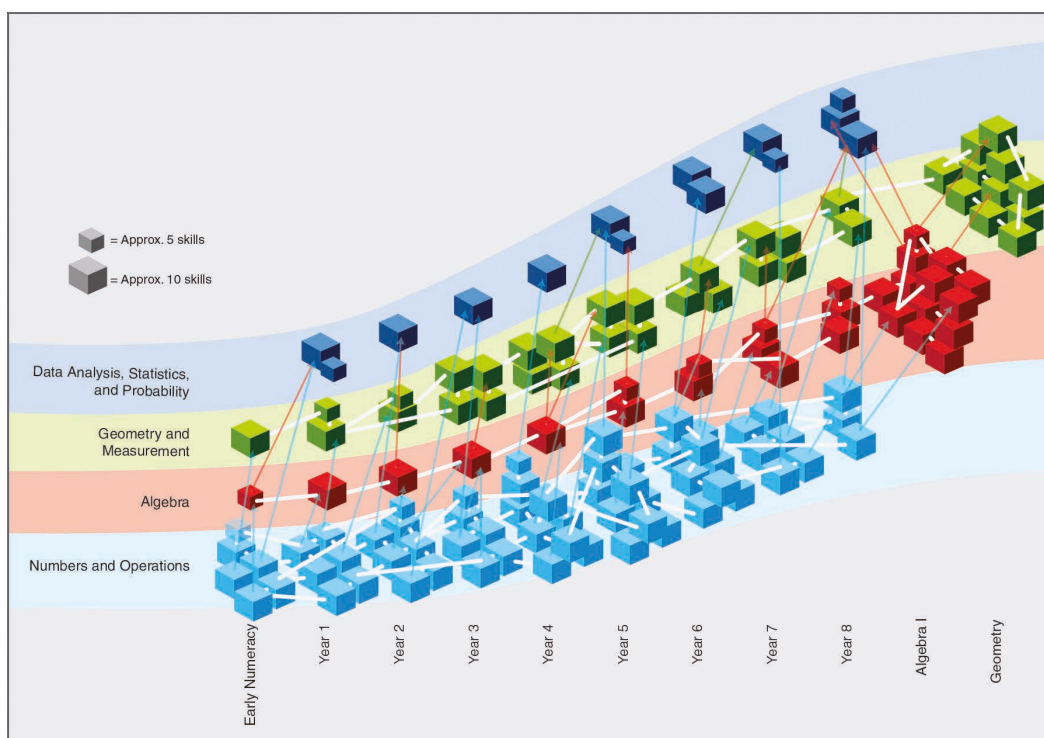


As a result of the statistical link between STAR assessments and Core Progress, a student's STAR assessment score provides insight into their achievement level, as well as skills and understandings they are ready to develop next. Learning Progressions are now an integral component of STAR forming a true bridge between assessment, teaching, and practice. They begin with early skills and progress to the levels of competence required to be higher education and career ready.

The skills and understanding in a learning progression provide the intermediate steps along with prerequisite skills necessary to reach the levels of expertise. These skills have been identified for the UK through the English National Curriculum. Renaissance Learning has worked with the National Foundation for Educational Research (NFER) to localise the Learning Progressions for both Maths and Reading so that schools in the UK can utilise STAR.

The learning progression is an interconnected web of prerequisite skills. Moving toward increased understanding over time requires continually building up and building on a solid foundation of knowledge, concepts, and skills. One indication of the interrelated network of concepts in Core Progress is the number of skills that build up and build on each other. A large proportion of core skills in the Learning Progressions serve as prerequisites to others in subsequent years.

Figure 2: Core Progress Path



The new National Curriculum provides a pathway to meeting educational targets. However, it does not describe a fully formed pathway along which students are expected to progress. Taking the National Curriculum as a starting point, NFER has developed a set of fully formed learning progressions which provide the intermediate steps and prerequisite skills necessary to reach the levels of expertise identified within it. The Learning Progressions begin with early skills and progress to the level of ability required for higher education and to be career-ready.

Framework for Learning Progressions - reading

The organisation of the Learning Progressions for reading covers four domains:

1. Word reading,
2. Comprehension,
3. Comprehension (literary/fiction/narrative),
4. Comprehension (information/non-fiction/non-narrative).

Organisation of National Curriculum Skill Areas within the four domains/eight sub-domains (headings)

Domain	Heading	Skill area	
Word reading	Phonic knowledge and skills	Grapheme-phoneme correspondences (GPCs)	Sounding and blending for decoding
	Word recognition	- Exception words (common and uncommon) - Root words, prefixes and suffixes	- Words with contractions - Syllables
	Fluency and accuracy	Reading aloud, pronunciation and speed	Automatic decoding, independent reading
Comprehension	Vocabulary	Vocabulary acquisition	
Comprehension (Literary/fiction/narrative)	Engaging and responding to texts	- Literary Heritage - Recital and Performance - Participation in discussion	Formulation and justification of opinion
	Understanding and interpreting texts	- Plot - Character - Setting - Summary - Main ideas and themes - Conventions - Author's purpose and perspective - Author's use of language	- Structure - Inference and evidence - Prediction - Comprehension monitoring - Purpose for reading - Discussion of understanding of texts - Analysis and comparison
Comprehension (Information/non-fiction/non-narrative)	Engaging and responding to texts	- Range of reading - Participation in discussion	Formulation and justification of opinion
	Understanding and interpreting texts	- Summary - Main ideas and themes - Conventions - Author's purpose and perspective - Author's use of language - Structure	- Inference and evidence - Comprehension monitoring - Purpose for reading - Research strategies - Discussion of understanding of texts - Analysis and comparison

Framework for Learning Progressions - maths

The organisation of the Learning Progressions for maths covers ten domains:

- | | | |
|--|--|--------------------------------------|
| 1. Number - number and place value | 4. Ratio, proportion and rates of change | 8. Geometry - position and direction |
| 2. Number - arithmetic operations | 5. Algebra | 9. Probability |
| 3. Number - fractions (including decimals and percentages) | 6. Measurement | 10. Statistics |
| | 7. Geometry - properties of shapes | |

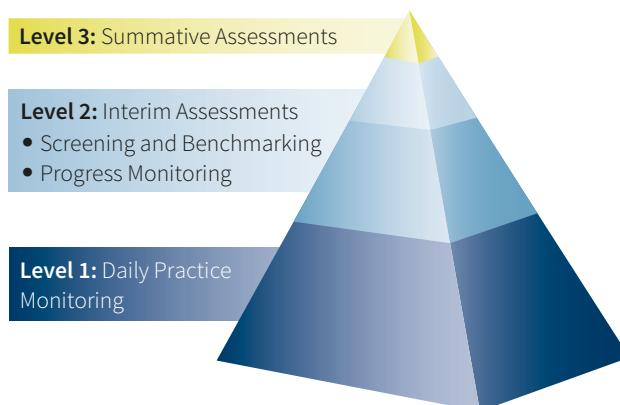
Organisation of National Curriculum Skill Areas within the ten domains

Domain	Skill area	
Number - number and place value	- Counting - Negative numbers - Reading, writing, ordering and comparing numbers - Number identification and representation	- Estimation, inverse operations, rounding and accuracy - Roman Numerals - Recognising place value - Problem solving with place value
Number - arithmetic operations	- Mental arithmetic methods - Written arithmetic methods - Problem solving using arithmetic	- Multiples, factors and prime numbers - Powers and roots - Standard Form
Number - fractions (including decimals and percentages)	- Decimals - place value - Decimals - relation of tenths, hundreds and thousandths - Decimals equivalents - Decimals - size order - Fractions - adding and subtracting	- Fractions - equivalency - Fractions - size order - Fractions - multiplication and division - Fractions - recognition - Percentage - Problem Solving
Ratio, proportion and rates of change	- Understanding and finding ratios - Using ratio in shapes and measurement	Interpreting graphic and algebraic representations of ratio
Algebra	- Expressions and equations - Functions	- Problem solving - mathematical contexts - Problem solving - real life scenarios
Measurement	- Key measurements - Conducting and recording measurements	- Solving problems involving measurements - Relationships and sequencing
Geometry - properties of shapes	- Identifying, describing and sorting geometric shapes - Problem-solving with geometric shapes	- Constructing geometric shapes - Circle theorem
Geometry - position and direction	- Translation, rotation, and coordinate geometry - Understanding & describing angles, segments & lines	- Problem-solving with angles, segments & lines - Constructing angles, segments & lines
Probability	Describing probabilities and using the probability scale	- Finding probabilities - Representing probabilities diagrammatically
Statistics	Data analysis: comparisons and mathematical relationships	- Data analysis: applying data - Data representation

The Renaissance Learning Information Pyramid

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

Figure 1: Renaissance Learning Information Pyramid



Level 1: daily 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 maths problem solving) or comprehension of direct teaching - both help to inform teaching and guide practice to improve student performance (e.g., Renaissance Learning's Accelerated Reader, Accelerated Maths, and MathsFacts in a Flash).

Level 2: interim assessments screening, benchmarking, and progress-monitoring assessments

Level 3: summative annual tests including once-a-year, high-stakes tests which assess student proficiency on national standards.

These three levels help to create a seamless, integrated three-stage assessment system. The system was pioneered by Renaissance Learning and reflects the model many experts and educational organisations recommend.

Renaissance Learning's interim assessments STAR Reading and STAR Maths make up the second, or middle, stage of the Renaissance Learning Information Pyramid. The purpose of interim assessments is to determine the extent to which teaching and learning tasks are strengthening students' abilities in key academic areas and preparing them to meet end-of-year proficiency targets. These assessments are administered regularly throughout the year to help determine how all students are progressing, both in groups and individually.

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. Screening and benchmarking periodic assessments, typically administered two to four times a year to monitor growth of a group towards a set target.
2. Progress-monitoring assessments, administered as often as 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 where support may be needed.

The value and cost of information

When choosing an appropriate educational assessment, it is important to have an assessment process that is both child and teacher friendly, minimizes lost teaching 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, and that can be purchased and supported within budgetary limits.

VALUE of an assessment	=	I Information —Amount of reliable & useful information produced by assessment
		C Cost —Total resources required, including price of acquisition; materials per administration; teacher time to administer, score, record, and interpret results; & time diverted from instruction

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 end up far more expensive in the long run. Two elements must be calculated:

- 1) the total cost of ownership
- 2) the value

Suppose an assessment is distributed for free but requires paper administration, necessitating the duplication of test instruments, scoring sheets, record sheets, and so on. 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 added to the calculation. A so-called one-minute test, in reality, may occupy as many as 10 minutes, on average, of the teacher's time per student per administration.

The total time considered must include preparing materials, explaining the assessment, the administration itself, recording and entering results, and the teacher's re-entry 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 teaching every time the test is administered - often weekly - multiplied by the number of measures that need to be taken.

This total cost, too, must be compared with 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 computerised 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.

Computer-adaptive testing

STAR Early Literacy, STAR Reading, and STAR Maths 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 and consistently more efficient than conventional tests, which present all students with the same test questions. 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"

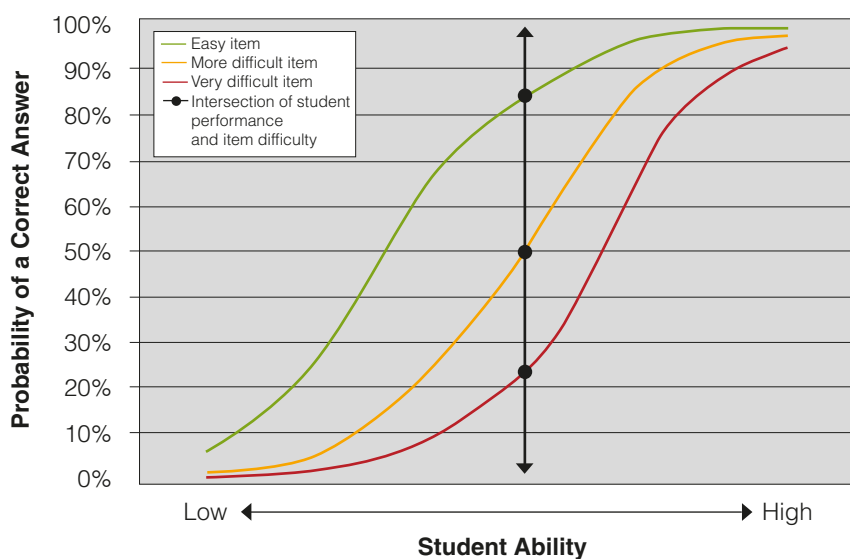
(Agdonini & Harris, 2010, p. 215).

The reliability and validity of the STAR assessments has been confirmed by key international groups including the National Foundation of Educational Research (NFER), the National Center on Response to Intervention and the National Center on Student Progress Monitoring, among others, 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 achieved in a number of 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, the probability of a correct response to an item can be calculated 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. The figure below shows the probability functions for three test items: an easy one, a moderately difficult one, and a still harder one.

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.

The figure below 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.

How computer-adaptive technology works

Item Development - Multiple-choice format

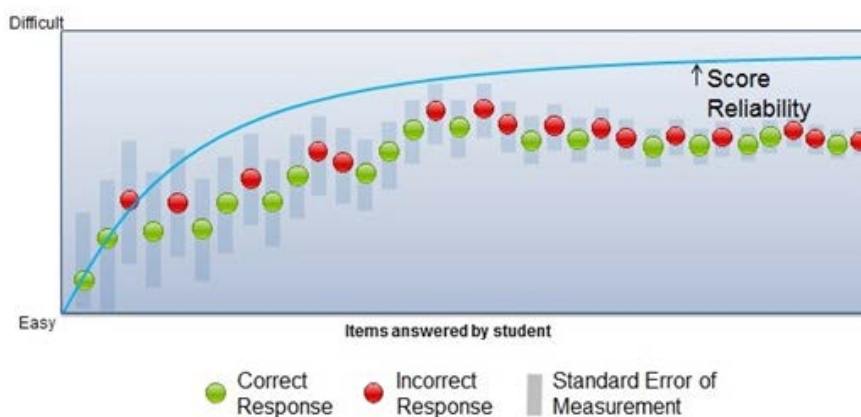
When the STAR assessments were developed, high priority was placed on selecting a test format that was well suited to computerised testing, precise, and efficient in terms of student and teacher time. Renaissance Learning explored, researched, discussed, and prototyped several item-response formats and 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):

Renaissance Learning constructs multiple-choice items to represent a balanced range of cognitive complexity. Item specifications require verifying the accuracy of all content; using year-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 computerised 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 & Armbusaiddi, 2000; Mattimore, 2009; Osterlind, 1998; Popham, 2003).

How Computer-Adaptive Technology Works



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 on-going item-writing and bias-and-fairness training, and adhere to the following process to ensure quality item development:

1. Analyse standards to be assessed in the categories of skill, action, vocabulary, and context; and refer to official resources for appropriate standard and year-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 dynamic calibration process.
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, patronising, 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 Maths and STAR Reading administer 34 items while STAR Early Literacy administers 27 items. The assessments were developed not only to provide precise measurement of student achievement in reading and maths, 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.

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

STAR Assessment Item Banks and Administration Breakdown, by Number and Type

	STAR Early Literacy	STAR Reading	STAR Maths
Number of items held	More than 2,400	More than 2,800	More than 1,900
Number of items	27 items	34 comprehensive items	34 items
Average Administration time	10 minutes	15 minutes	20 minutes

Dynamic Calibration

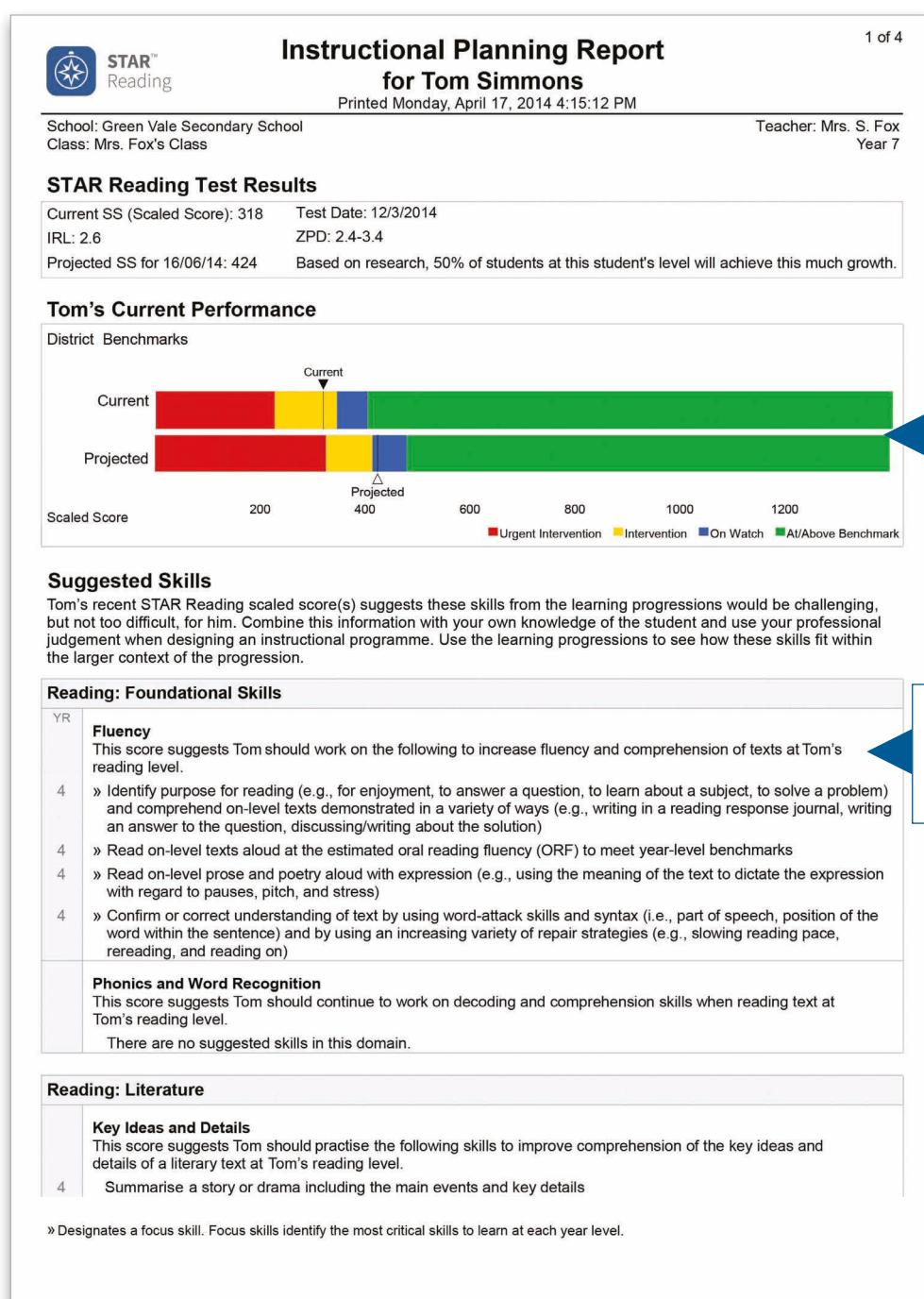
To maintain and update the large item banks for each STAR assessment, Renaissance Learning continually develops and calibrates new test items using a special feature called dynamic calibration. In dynamic calibration, one or more new items are embedded at random points in a STAR test. These items do 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.

Instructional planning

After a student takes a STAR assessment, the software places the resulting Scaled score on the Learning Progression, which reports skills the student has likely mastered, those they are ready to develop next, and suggestions for the teacher to focus instruction. In essence, Learning Progressions serve as a road map to help teachers understand both where students have been and where they need to go to become university and career ready.

Instructional Planning Reports from STAR provide teachers lists of skills individual students – and at the class level, groups of students – are ready to develop next (see student example below). With Learning Progressions, teachers can search for the skills and domains listed on the reports to further focus next steps for students.

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



Growth Modelling

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

With STAR Early Literacy, STAR Reading, and STAR Maths, educators have access to a scientific method for setting appropriate, achievable, and challenging progress-monitoring targets for students. Since thousands of schools use the STAR assessments, Renaissance Learning is able to observe how students develop skills over time.

Using longitudinal data on the learning patterns of more than 1 million students for reading, and nearly 350,000 students for maths, the STAR assessments provide educators with critical information about how students grow over time. Specifically, the Target-Setting Wizard in each STAR assessment uses this information to help educators set progress-monitoring targets personalised to each student - targets that are challenging but reasonable.

STAR Reading ADMINISTRATOR USER, School Network 2013 - 2014

Home > Screening, Progress Monitoring & Intervention > Student Detail > Manage Targets [Manuals](#) [Help](#) [Log Out](#)

Manage Targets
Define an intervention and set a target

School: **Renaissance Learning Academy**
Student: **Martin, Andrea**

Latest Test	Score	Target	Growth Rate
25/10/2013	148 SS / 2 PR	--	calculated after four scores

Intervention Details **Interpretation & Guidance**

Intervention Name
Appears in report details: Intervention 1

Target End Date
Used for SS/week calculation: 30/06/2014

Target
Expected growth rate and score

Starting test: 25/10/2013 - 148 SS / 2 PR
(Sets intervention line; starts trend and target lines)

Reference points to help you select a target type:
 - Maintain 2 PR throughout the school year = 0.0 SS/week
 - Reach 50 PR benchmark by the end of the school year = 13.6 SS/week

Select a target type (based on students who scored similarly*)

☒ **Moderate:** 2.0 SS/week = (press Calculate Target button)

☐ **Ambitious:** 4.2 SS/week = (press Calculate Target button)

Or define a custom target:

☐ Growth Rate

*Research data shows that 50% of students who started at the 2 PR were able to achieve a Moderate growth rate or better, while 25% were able to achieve an Ambitious growth rate or better. Set an appropriate target for this student and adjust as necessary.

[Cancel](#) [Calculate Target](#) [Save](#)

© 2014 Renaissance Learning, Inc. All rights reserved. [About Renaissance Place](#)

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 with students who score in the same decile, making the Target-Setting Wizard growth predictions much more accurate than a “one-size-fits-all” growth rate. Using growth modelling data, the Target-Setting Wizard offers research-based progress monitoring recommendations called “Moderate” and “Ambitious” targets. A moderate target is a growth target that 50% of students nationally with the same starting score would reach. Ambitious targets 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 Target-Setting Wizard, professional judgment can now be informed by research.

After a student has taken an initial STAR assessment and the teacher has selected a target for that student, a target line appears on the STAR Student Progress Monitoring Report. The target line depicts the rate of growth the student must attain to meet the selected target. Following subsequent STAR tests, a trend line showing the student's actual growth rate is automatically drawn on the report.



Student Progress Monitoring Report

1 of 2

Wednesday, 19 December 2012 14:29:18

School: Renaissance Academy

Reporting Period: 1/8/2012-31/7/2013
(School Year)

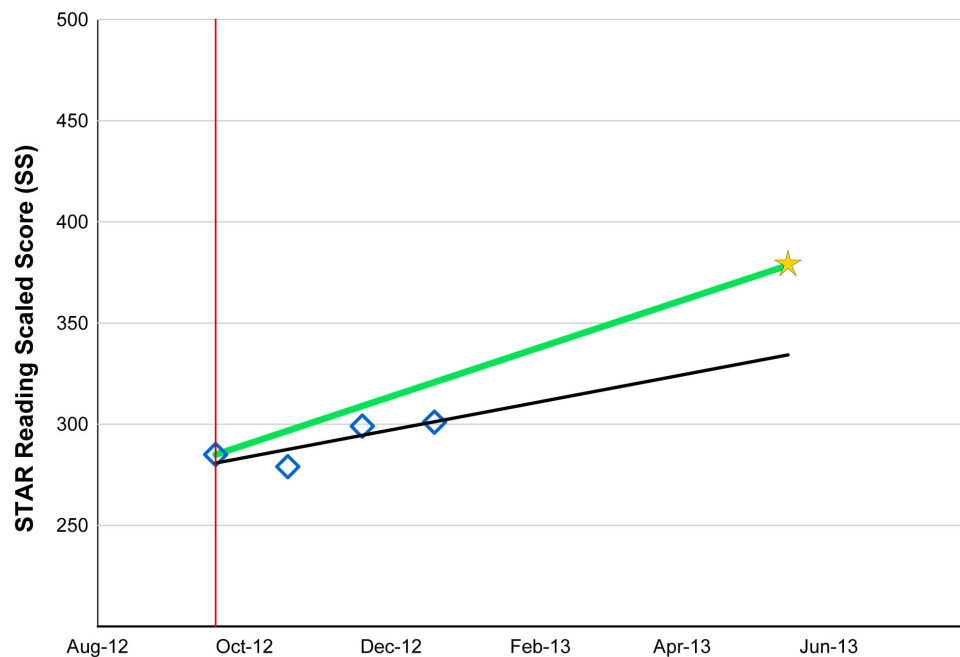
Gough, Catherine

Year: 5

Class: Y5AR

ID:.....

Teacher: Simmons, L



◇ Test score

— Trend line is statistically calculated after four or more tests to show the direction the scores are moving.

— Target line represents the student's expected growth path toward the target.

★ Star represents the student's current target.

| Intervention line identifies the start date of an intervention programme.

Catherine's Current Target

Target : 379 SS / 22 PR (Moderate)

Expected Growth Rate: 2.8 SS/Week

Target End Date: 15/5/2013

Actual Growth Rate: 1.6 SS/Week

Fluctuation of scores is typical with any test administered multiple times within a short period. Focus on the general direction emerging after multiple administrations of the test rather than on the ups and downs between individual scores. The student's trend line and growth rate appear on the report after taking four tests.

R56477



Student Progress Monitoring Report

1 of 2

Thursday, January 23, 2014 3:22:25 PM

School: Lake View School

Reporting Period: 03/9/2013-23/1/2014
(Term 1)

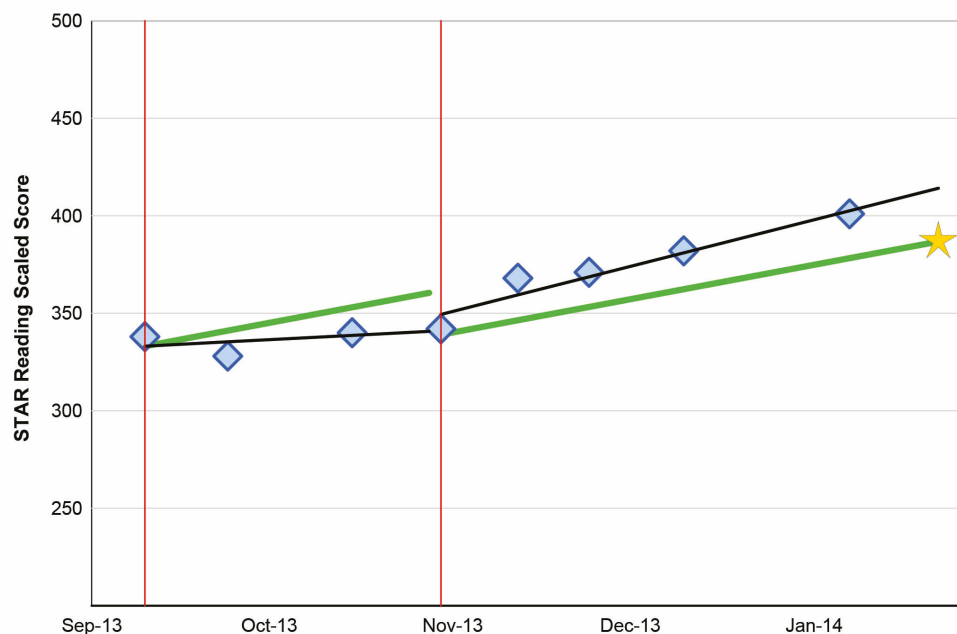
Taylor, Mia

Year: 5

Class: G5 (Davis)

ID: TAYLORM

Teacher: Davis, J.



Enterprise Test



Trend line is statistically calculated after four or more tests to show the direction the scores are moving.



Target line represents the student's expected growth path toward the goal.



Star represents the student's current target.



Intervention line identifies the start date of an intervention program.

Mia's Current Goal

Goal: 387 SS / 14 PR (Ambitious)

Goal End Date: 23/1/2014

Expected Growth Rate: 4.0 SS/Week

Fluctuation of scores is typical with any test administered multiple times within a short period. Focus on the general direction emerging after multiple administrations of the test rather than on the ups and downs between individual scores.

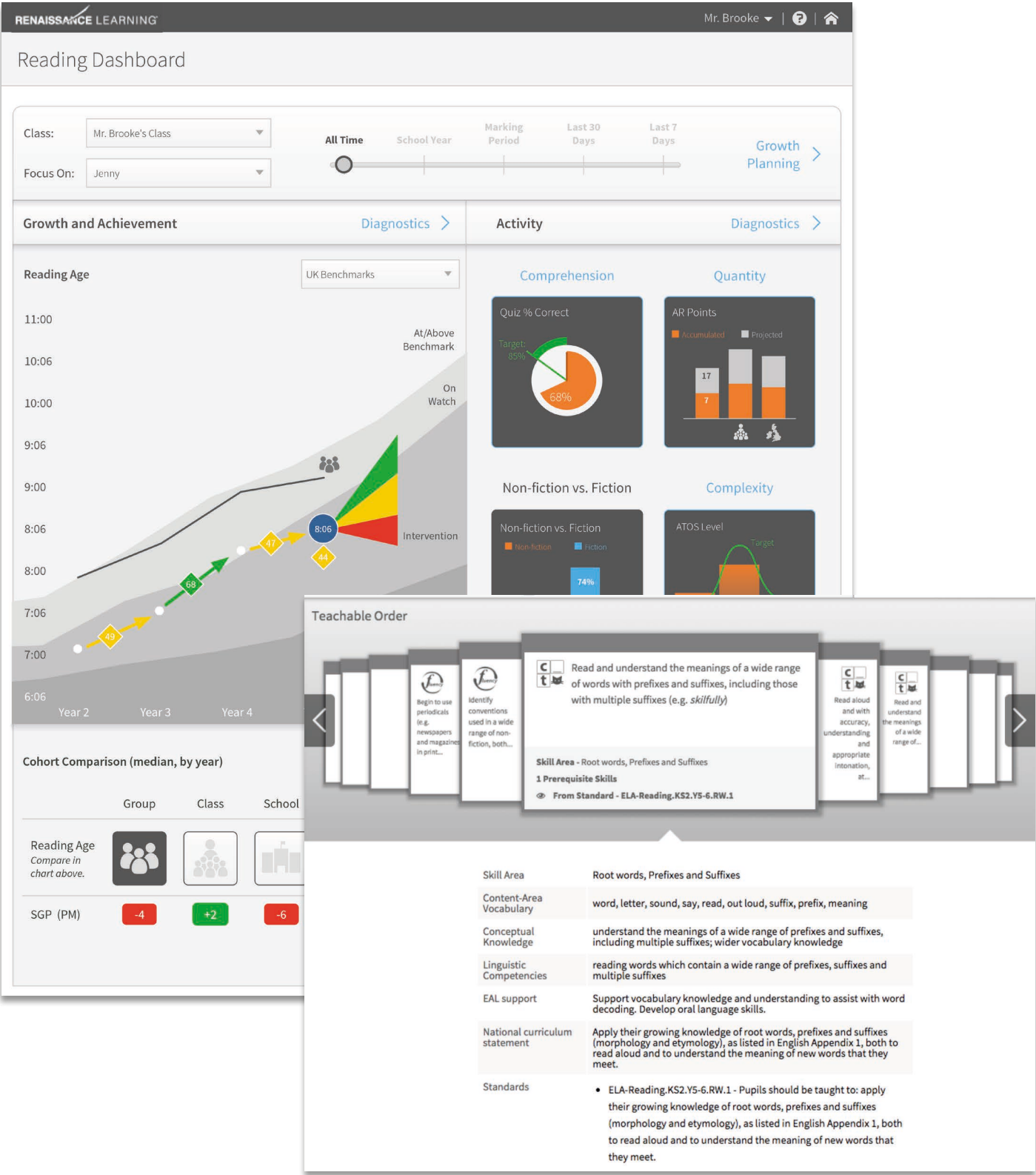
By comparing the target and trend lines, educators can determine whether a student's growth trajectory is steep enough for the student to reach the target. Educators can then use this information to make the best teaching support decisions. The breadth and depth of our database allows us to identify 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 targets to help students reach their full potential.

Reading and Maths Dashboards

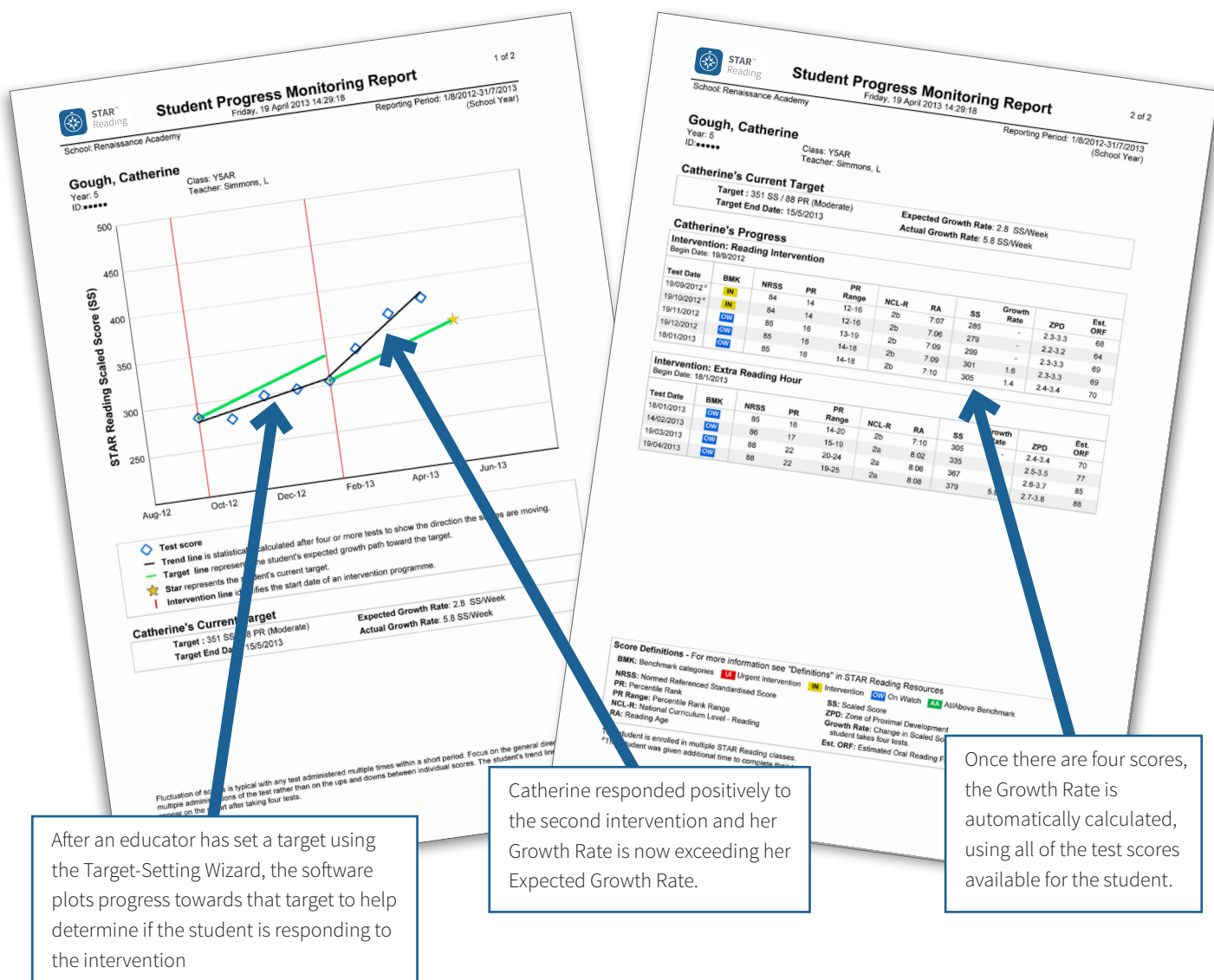
The Dashboard offers a streamlined way for teachers to plan for individuals or groups of students, by providing details on students' current performance, projected growth and a suggested focus for new skills.

The suggested skills screen in the Dashboard not only provides a way to view the prerequisite skills, but also shows the next skills in the progression. Teachers will be able to reference the year-level domain expectations to help see how the discrete skills will move students through the different levels of understanding.



A closer look at the STAR assessments

The STAR assessments allow teachers to precisely and efficiently assess student achievement in pre-reading skills (STAR Early Literacy), reading (STAR Reading), and maths (STAR Maths). Teachers use the wealth of data provided by the assessments to target teaching, provide students with the most appropriate teaching materials, and intervene with struggling students. Teachers access STAR assessment data via informative reports. For additional information, full technical manuals are available for each STAR assessment.



Another way to view data

In addition to the reports available in STAR Early Literacy and STAR Reading, Renaissance Learning has developed the STAR Learning to Read Dashboard. Teachers can zero in on their emergent readers' progress and view the percentage of students classified as Probable Readers with at least one STAR Early Literacy or STAR Reading test taken in the school year to date.



About the STAR Early Literacy assessment

The STAR Early Literacy assessment is a reliable 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 teaching to meet 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 skill domains assessed in STAR Early Literacy, including general readiness, graphonomemical knowledge and structural analysis.

The STAR Early Literacy domains include four of the five critical areas of reading instruction. 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 Oral Reading Fluency score for beginning readers. Oral reading fluency is the number of words a student should be able to read correctly on a year-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 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.

Early Literacy Item

This item measures: Sound-Symbol Correspondence: Consonants

Early Numeracy Item

This item measures: Composing and Decomposing

Using STAR Early Literacy data

STAR Early Literacy is used for screening/benchmarking and progress monitoring of emergent readers. The assessment also provides diagnostic data to make teaching decisions and help identify likely gaps in knowledge for students experiencing reading difficulties.

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 teaching and curriculum frameworks. Its quick and accurate results provide teachers with specific benchmarking, screening, and progress-monitoring information to help tailor teaching, monitor reading growth, and improve reading achievement for all students. STAR Reading assesses reading comprehension through the use of short comprehension questions and skill based questions. Their use is based on abundant and long-standing research verifying that vocabulary is closely tied to comprehension.

The information needed to determine the correct answer is given within the assessment question 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 and answer choices. The questions require reading comprehension because the student must actually interpret the meaning of the sentence 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 across all the school years.

STAR Reading results for students in lower years include an Oral Reading Fluency score. 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. It automates benchmarks, cut scores, progress-monitoring targets, and teaching 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 Zone of Proximal Development (ZPD), or individualised reading range, for use within Accelerated Reader™ (AR™). To experience optimal growth, the student chooses books with readability levels within this range.

About the STAR Maths assessment

The STAR Maths assessment is a reliable, valid, and time-efficient assessment of mathematics skills appropriate for use within various teaching and curriculum frameworks. Its quick and accurate results provide teachers with specific benchmarking, screening, progress-monitoring, and diagnostic information to help tailor teaching, monitor maths growth, and improve maths achievement for all students. The content for STAR Maths is based on analysis of national standards, various curriculum materials, test frameworks, and content-area research, including best practices for mathematics teaching. Research indicates that numeration concepts are key for deep conceptual development and that computational processes emphasizing fluency complement conceptual development. STAR Maths 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 Maths item bank includes core maths objectives, with multiple items available to measure each objective.

Using STAR Maths data

STAR Maths is used for screening/benchmarking and progress monitoring of students. It automates benchmarks, cut scores, progress-monitoring targets, diagnosis of students’ skills, and teaching recommendations, and helps the teacher determine if student achievement is heading in the right direction.

Purpose and frequency

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

Reliability and Validity of the STAR assessments

In 2009 STAR Reading and STAR Maths were among the first assessments to be highly rated in the USA by the National Center on Response to Intervention (NCRTI) for screening and progress monitoring. In subsequent reviews, the STAR assessments have maintained the NCRTI's highest ratings.

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. In the UK an equating study was carried out in autumn 2006 by the National Foundation for Educational Research (NFER) on behalf of Renaissance Learning, to provide validation evidence for the use of the Renaissance STAR Reading and STAR Maths tests in English schools. The report concluded that the strong correlations provide evidence that both STAR Reading and STAR Maths are suitable for use in England.

A copy of the full report can be downloaded from here: <http://eric.ed.gov/?id=ED502442>

The following provides a brief explanation of the reliability and validity of each STAR 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 favourably 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 year. 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 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.

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 enrolment, and socioeconomic status. The reliability estimates were very high, comparing favourably 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 correlations were substantial and reflect well on the validity of STAR Reading as a tool for assessing reading achievement. 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 re-normed, 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.

Since then STAR Reading has been localised for use in UK schools, and normed to national standards in the UK using a sample of 816,429 tests by Dundee and Stirling Universities. High levels of correlation with tests such as the Suffolk Reading Scale and teacher assessments have shown its reliability and validity for schools in the UK.

STAR Maths reliability and validity

STAR Maths 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 enrolment, and socioeconomic status. The final norming sample for STAR Maths included approximately 29,200 students from 312 schools in 48 U.S. states. The reliability estimates were very high, comparing favourably with reliability estimates typical of other published maths achievement tests.

For STAR Maths to measure maths achievement, Renaissance Learning knew it was necessary that its scores correlate highly with other measures of maths achievement. STAR Maths has been localised for use in UK schools, and normed to national standards in the UK using a sample of 24,334 tests, from children of all ages across the UK. The resulting correlation estimates were substantial and reflect well on the validity of STAR Maths as a tool for assessing maths achievement. As with STAR Reading, additional data supporting the validity of STAR Maths 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. High levels of correlation have indicated its use and reliability and validity for schools in the UK.

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 efficient 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 colours, 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 favourite 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 recognising 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 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, recognising the alphabet, naming letters, recognising 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; recognising 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 per cent of mastery of specific skills within the seven domains listed above.

STAR Reading

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 Reading scaled scores range from 0 to 1400. All norm-referenced scores are derived from the scaled score.

Zone of Proximal Development (ZPD) the range of difficulty level of books the student should read to allow for independent reading. 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.

National Curriculum Reading Level - English National Curriculum Level including sublevels.

Reading Age (RA) is correlated, by NFER, to the Suffolk Reading Test. This has proved to have a high correlation, so STAR Reading can give a good indication of RA.

Norm Referenced Standardised Scores (NRSS) which help you to see how a student compares nationally with others of a similar age. A score of 100 is average, higher than 100 indicates the student is above average and below 100, below average.

Percentile rank (PR) is a norm-referenced score that provides a measure of a student's score compared with other students of the same age nationally. The percentile rank score, which ranges from 1 to 99, indicates the percentage of other students of a similar age nationally who obtained scores equal to or lower than the score of a particular student.

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.

Oral Reading Fluency (ORF) is provided for students between Years 2 and 5. ORF is a calculation of a pupil's ability to read words quickly and accurately in order to comprehend text efficiently.

STAR Maths

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 Maths scaled scores range from 0 to 1400. All norm-referenced scores are derived from the scaled score.

Accelerated Maths Library Recommendation - which Year Group Maths Library is most suited to the student based on the results of their STAR Maths assessment. This helps educators place a student in the Accelerated Maths library that will be of the most benefit, based on that student's individual achievement level.

National Curriculum Maths Level - English National Curriculum Levels including sublevels.

Norm Referenced Standardised Scores (NRSS) which help you to see how a student compares nationally with others of a similar age. These scores will show schools how students compare with a similar age group of students across the country.

Percentile rank (PR) is a norm-referenced score that provides a measure of a student's score compared with other students of the same age nationally. The percentile rank score, which ranges from 1 to 99, indicates the percentage of other students of a similar age nationally who obtained scores equal to or lower than the score of a particular student.

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 Maths assessment is taken again within a short time frame - for example, 4 to 6 weeks.

Conclusions

The aim of schools is to ensure that all students are fully prepared for a continued education or career. The benefit of STAR assessments is that they lay out a pathway to guide teaching and learning over time so that student competence in the domain can be advanced coherently and continuously. Learning Progressions for Maths and Reading are at the heart of this process. STAR helps teachers locate where students are on their pathway, not only pointing in the right direction, but also providing tangible and achievable next steps for getting there.

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*
- 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.

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 Maths

- 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.



STAR Early Literacy, STAR Maths, STAR Reading, STAR and Renaissance Learning are trademarks of are trademarks of Renaissance Learning, Inc. and its subsidiaries, registered, common law, or pending registration in the United Kingdom, United States and other countries. All other product and company names should be considered as the property of their respective companies and organisations.