New Thinking in Response to Intervention
A Comparison of Computer-Adaptive Tests and Curriculum-Based Measurement Within RTI

Edward S. Shapiro
Center for Promoting Research to Practice
Lehigh University
Reports and software screens may vary from those shown as enhancements are made.

©Copyright 2012 by Renaissance Learning, Inc. All rights reserved. Printed in the United States of America. All logos, designs, and brand names for Renaissance’s products and services, including but not limited to Accelerated Reader, Accelerated Reader 360, Core Progress, Star Early Literacy, Star Custom, Star Math, Star Reading, Renaissance, and Renaissance Learning, are trademarks of Renaissance Learning, Inc., and its subsidiaries, registered, common law, or pending registration in the United States. All other product and company names should be considered the property of their respective companies and organizations.

This paper was commissioned for production by Renaissance Learning but represents solely the opinion of the author. Correspondence concerning this paper should be addressed to Edward S. Shapiro, Ph.D., Director, Center for Promoting Research to Practice, Lehigh University, L-111 Iacocca Hall, 111 Research Dr., Bethlehem, PA 18015. Email: ed.shapiro@lehigh.edu.

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
P.O. Box 8036
Wisconsin Rapids, WI 54495-8036
(800) 338-4204
www.renaissance.com
educatordevelopment@renaissance.com

9/18
Contents
Introduction .................................................................................................................................................................................... 6
Response to Intervention Defined .................................................................................................................................................. 7
Comparison of CAT and CBM ...................................................................................................................................................... 10
Discussion .................................................................................................................................................................................... 19
Conclusion .................................................................................................................................................................................. 20
References ................................................................................................................................................................................ 21
Acknowledgements ..................................................................................................................................................................... 22

Figures
Figure 1: RTI Pyramid ................................................................................................................................................................. 6
Figure 2: Star Screening Reports .................................................................................................................................................. 12
Figure 3: CBM Screening Report ............................................................................................................................................... 13
Figure 4: CBM Screening Reports ............................................................................................................................................... 13
Figure 5: CBM Progress-Monitoring Report ............................................................................................................................... 14
Figure 6: Star Goal Setting Wizard ............................................................................................................................................ 15
Figure 7: Star Progress-Monitoring Report ................................................................................................................................ 16
Figure 8: Understanding Star Scaled Score .................................................................................................................................... 17
Figure 9: Star Diagnostic Report ............................................................................................................................................... 18
Figure 10: Star Instructional Planning Report ........................................................................................................................... 18

Table
Table 1: Components of RTI ......................................................................................................................................................... 8
Dear Educator,

When I was asked to consider writing this practice brief, I was, at the time, consulting with a number of elementary schools in the planning and implementation of Response to Intervention (RTI) models.

One of the first questions tackled by those schools was what measure should be selected for universal screening and progress monitoring. The research literature as well as my own experience immediately turned to Curriculum-Based Measurement (CBM) as a logical recommendation for these schools. When the schools examined the nature of CBMs, especially for reading and early literacy, many asked whether these were the only measures available, and if there were other options that shared equally strong research support. In consulting the National Center on Response to Intervention Screening and Progress Monitoring Tools Charts, it was evident that there were many other choices.

CBM has served those implementing RTI models very well for many years, especially RTI models focused on reading and early literacy. Indeed, the research support for CBM within RTI models remains very strong. It is likely that CBM will continue to be used extensively in RTI models in the future.

However, when we look at the national investment of the U.S. Department of Education that is currently underway in the development of next-generation statewide assessment measures, Computer-Adaptive Testing is emerging as the foundation of the construction of these tests. As RTI models evolve, it is likely that educators will be looking to Computer-Adaptive Testing for universal screening and progress monitoring to enhance more immediate instructional decision-making processes called for within those models.

My purpose in writing this practice brief is to show educators that while CBM has and continues to serve us well as key measures for RTI models, there are alternatives to CBM that can provide the support needed to make informed instructional decisions.

The Star measures, a set of Computer-Adaptive Tests for reading and math, are one option that can help move RTI into next-generation testing using the assessment measures of the future, today.

Edward S. Shapiro, Ph.D.
October 14, 2011
Introduction

Over the past few years, there has been increasing interest in the process known as Response to Intervention (RTI) within the education community. RTI is a multi-tiered system of supports that organizes services to all students in an effort to intervene early with those students whose performance is predictive of subsequent failure. Applied to both academics and behavior, RTI is increasingly viewed as an important and major effort to reform the delivery of instruction, especially at the elementary level.

Key to the RTI process is the use of assessment tools that provide universal screening and progress monitoring. These assessments work together to help educators identify students with trajectories of learning that are not likely to lead to grade-level goals. They also provide formative feedback about the success of interventions used to shift a student’s trajectory to meet grade-level goals.

Among the many methods of assessment, Curriculum-Based Measurement (CBM) has been most closely associated with RTI models. CBM has a long and well-established research base that makes it a logical and natural choice. However, equating CBM with RTI is unfortunate as it is not the only assessment process that meets the requirements of an effective RTI system. The purpose of this paper is to compare two research-based systems of assessment useful within an RTI model: Computer-Adaptive Testing (CAT) and Curriculum-Based Measurement.

The paper begins with a brief examination of RTI, presents the key components of RTI, and details the assessments used within the RTI model. Examples that show both the CAT and CBM approaches are presented, with a final brief discussion contrasting the two systems.

Figure 1. RTI Pyramid
Response to Intervention Defined

What we do in the early years of formal schooling makes a huge difference in the academic health of a child’s development. In the earliest years of school, children must learn to read, do math, and write. We know that students who struggle in first grade, especially in reading, have a very high probability of academic struggles throughout school. In one of the most widely cited studies, Juel (1988) showed that children who fail to master basic reading skills at the end of first grade have an 88% probability of still showing struggles in fourth grade. Many others over the past two decades have replicated these findings (e.g., Abbott, Berninger, & Fayol, 2010; Silberglitt, Burns, Madyun, & Lail, 2006; Snow, Porche, Tabors, & Harris, 2007).

As a part of the reauthorization of the Individuals with Disabilities Education Act (IDEA), Congress recognized the importance of the early identification of children who are potentially at risk of academic failure. IDEA allows schools to use a process for identifying students with specific learning disabilities (SLD) that examines the degree to which students respond to high quality, research-based interventions. The provision is considered a landmark shift in the conceptualization of SLD. It enables educators to use a process that differentiates between students who respond to instruction designed to shift their learning trajectory, from those who need much more intensive forms of specialized instruction to sustain academic success (i.e., those in need of special education). The term “Response to Intervention” emerged as the way to describe this process.

Although RTI began as a way to assist educators in determining the presence of a specific learning disability, it quickly emerged as much more. RTI is viewed as a method for reforming the way academic services are delivered to all students.

RTI is often conceptualized as a triangle (Figure 1). The instructional practices at the base of the triangle, known as core instruction or Tier 1, are the foundation on which all future academic success rests. When the Tier 1 instruction is research based and delivered accurately, most students achieve academic success without additional intervention.

Despite strong core instruction, some students show evidence that if their learning trajectory continues at the current rate, they will not meet the expected academic goals. These students need additional instructional supports supplemental to core instruction. Typically called Strategic or Tier 2 support, the instructional interventions at this tier add a dimension of teaching to address the identified needs of students. The objective is to correct a student’s trajectory of learning and erase the gap between targeted students and peers. This is done by focusing on three key components: small-group instruction, a specified number of minutes per day/week, and teachers trained to deliver the remedial instruction.

Students who do not respond to Tiers 1 and 2 are placed in an intensive, or Tier 3 level of support, which is supplemental to Tier 1. Students in Tier 3 were not responsive to Tier 2 interventions or were identified earlier (based on universal screening measures) as in need of immediate intensive interventions. Students identified as needing Tier 3 support are at high risk for future academic failure. They need a level of support that requires focused instruction, very small groups, and extensive intervention. This is best accomplished by school personnel who have expertise working with students requiring more intensive intervention. Students at Tier 3 who are found to not respond to interventions at rates that would significantly alter their learning trajectory are considered as potentially eligible for determination as students with a specific learning disability.
Components of RTI

An RTI process consists of key components, including universal screening, high-quality core instruction, progress monitoring, tiered interventions, collaborative data-based decision making, parent involvement, and administrative support (Table 1). A full description, discussion, and details of these components is beyond the scope of this paper and can be found in excellent practitioner resources such as Burns and Gibbons (2008), Quinn (2009), and Wright (2007). This paper focuses on assessment within RTI.

<table>
<thead>
<tr>
<th>Table 1. Components of RTI</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Universal Screening</strong></td>
</tr>
<tr>
<td>• All children assessed at beginning, middle, and end of school year on skills identified as highly predictive of future success or failure</td>
</tr>
<tr>
<td>• Assesses the overall success of the core instruction (Tier 1)</td>
</tr>
<tr>
<td><strong>High-Quality, Standards-Aligned Instruction</strong></td>
</tr>
<tr>
<td>• Core instruction delivered to all students using a research-based, empirically supported program that is closely aligned to state standards and/or Common Core State Standards</td>
</tr>
<tr>
<td><strong>Progress Monitoring</strong></td>
</tr>
<tr>
<td>• Conducted on an ongoing basis over time for students who are in need of tiered supports beyond Tier 1</td>
</tr>
<tr>
<td>• Assessment frequency is more periodic than universal screening, usually at least once per week</td>
</tr>
<tr>
<td>• Data are sensitive to student improvement over time; sufficient data to establish a reliable trend must be collected</td>
</tr>
<tr>
<td><strong>Tiered Interventions</strong></td>
</tr>
<tr>
<td>• Supplemental, research-based instructional interventions to core instruction derived from a problem-solving process and focused on student needs</td>
</tr>
<tr>
<td>• Usually delivered in small groups, with larger groups at Tier 2 than Tier 3</td>
</tr>
<tr>
<td><strong>Collaborative, Data-Based Decision Making</strong></td>
</tr>
<tr>
<td>• Teams of school professionals examine multiple data sources to discuss the appropriate intervention to impact the child</td>
</tr>
<tr>
<td>• The number and nature of the team structure is often defined by the local context (i.e., school-level versus grade-level data teams)</td>
</tr>
<tr>
<td><strong>Parental Involvement</strong></td>
</tr>
<tr>
<td>• Engagement of parents in the process of understanding and supporting the efforts to provide instructional support</td>
</tr>
<tr>
<td>• Maintain close and frequent communication with parents about student progress</td>
</tr>
<tr>
<td><strong>Administrative Support</strong></td>
</tr>
<tr>
<td>• Leadership at central, building, and teacher levels provide key supports to the process</td>
</tr>
<tr>
<td>• Administrative support for infrastructure, schedule, materials, ongoing professional development, and building consensus</td>
</tr>
</tbody>
</table>
Assessment Processes within RTI

RTI is only successful when all the components listed in Table 1 are simultaneously in place. However, assessment plays a pivotal role. The entire RTI process relies on the accurate and effective use of assessment for universal screening and progress monitoring. These two processes provide the focus for decisions made by collaborative teams and directions for instructional changes needed to improve student performance.

*Universal screening* provides periodic windows into student performance by comparing against the performance of peers. The measures are obtained at intervals across the school year (usually beginning, middle, and end of year) and used to establish the expected level of growth of typical performing peers. Benchmark assessments should be relatively brief, inexpensive, easily administered, and easily scored. With screening, it is understood that more students will be identified as at risk than truly are. Likewise, there will be some students not identified who are found to be having difficulties. As such, the benchmark measure must be combined with additional data in making important diagnostic decisions about student performance.

*Progress monitoring* plays a critical role in evaluating how well a student is responding to intervention. Progress-monitoring measures must be frequent, sensitive to instructional change over a short period of time, predictive of overall success as measured by the benchmark assessment, and able to drive instructional decisions. Progress-monitoring measures must assist educators in determining if interventions are effective and, perhaps more importantly, what to do in cases when they are not.

Progress-monitoring measures must assist educators in determining if interventions are effective and, perhaps more importantly, what to do in cases when they are not.
Comparison of CAT and CBM

Curriculum-Based Measurement

Probably the most used and accepted assessment tool to date within RTI is Curriculum-Based Measurement (CBM). CBM measures are rate based, efficient, and easily administered. The reading portion is administered to students individually in approximately 1-2 minutes per child. The math section is given to students in small groups and takes 5-10 minutes per group. As a rate-based measure, fluency (i.e., accurate responses per unit of time) is the key outcome of student performance.

First developed by Deno and his colleagues (e.g., Deno, 1985, Deno, Marston, & Tindal, 1985), CBM was designed to serve as an index of overall growth and skill development across curriculum objectives. The measures were viewed as single indices to signal change in student learning in the area being assessed. For example, a measure of the oral reading fluency (words read correctly within one minute) was found to be a very strong indicator of overall acquisition of reading skills (Deno, Mirkin, & Chiang, 1982). It reflects a student’s overall performance in the skills embedded within learning to read: phonemic awareness, phonics, fluency, vocabulary development, and comprehension. A student’s score on an Oral Reading Fluency (ORF) measure is then used to index how well reading instruction is resulting in improved reading performance. Substantial and repeated research has shown that measures of oral reading fluency are highly predictive and related to all skills embedded in reading performance, even well into middle school (Denton et al., 2011; Petscher & Kim, 2011), although the correlations to comprehension reduces significantly after fourth grade (Shapiro, Solari, & Petscher, 2008). In addition, CBM measures in reading have been found to be highly predictive of outcomes on state assessments (e.g., McGlinchey & Hixson, 2004; Shapiro, Keller, Edwards, Lutz, & Hintze, 2006; Silberglipt, Burns, Madyun, & Lail, 2006).

It is important to understand that while CBMs index overall growth, they are not demonstrations of specific skills. Similarly, while CBM reading indexes whether the instructional interventions are having the desired impact, one would not directly teach children to read out loud faster as a way to increase performance.

In other words, while fluency is what we measure, increasing fluency in reading is not necessarily the skill that needs to be targeted. Many educators using CBM measures make this critical mistake and believe that because we measure a student’s fluency to gauge overall reading performance, a student’s fluency must then be the target for improvement. Although fluency may be an appropriate concern for improving a student’s reading skills, one would not make that determination on the CBM metric alone. CBM measures tell the educator that a student is struggling in reading, but do not point specifically at the needed skills for intervention.

Common Misconceptions

A common misconception in RTI is that the skills taught in intervention must be assessed separately. On the contrary, there are several ways to monitor a student’s progress in a skill area, including computer-adaptive measures that assess a broad range of skills with each test. In other words, it is possible to get a good indication of student progress in phonics with a test that assesses phonemic awareness, phonics, vocabulary, and comprehension at one time.

Common Misconceptions

It is a common misconception that CBMs assess specific skills. For example, it is often suggested that a Nonsense Word Fluency (NWF) probe assesses phonics. This is not the case. NWF assesses a student’s overall ability to decode but does not identify which phonics skills a student may lack. Is the student struggling in initial consonant sounds? Is the student struggling with medial vowels? A student’s NWF score doesn’t answer these questions. One would never design an intervention to teach students to read nonsense words aloud faster. Therefore, determining interventions based on the NWF score requires additional diagnostic information.
**Computer-Adaptive Testing**

Recently, Computer-Adaptive Testing (CAT) has emerged as an important option for assessment within RTI models. Based on the Item Response Theory (IRT) approach to test construction, CATs adjust the items administered based on student responses and the difficulty of the items. In other words, when a student answers a question correctly, the student is then given a more difficult question. Student responses cue shifts in subsequent items and result in an indication of the skills attained and not attained across an assessment domain.

These assessments include large item banks, are not timed, and are based on accuracy of student responses. Because CAT measures are carefully calibrated, the test quickly pinpoints the skill sets that represent a student’s academic achievement level. Some CAT assessments can be given in 10 to 25 minutes and are individually administered via computer.

It is important to understand that Computer-Adaptive Tests adjust based on a student's accuracy in answering items, regardless of the time needed to answer the question. In contrast to CBM measures that use fluency as the key indicator of student performance, CATs use a student's actual level of right and wrong answers to select the next question from a vast item bank that spans multiple domains. As such, CAT measures are much more focused on skills within various domains and sub-domains of academic areas compared to CBM measures.

Although CBM has been widely accepted and adopted within RTI models, other assessments also meet the defining characteristics of screening and progress monitoring—including some Computer-Adaptive Tests. These measures can serve as universal screening tools. They report on the relative standing of students compared to their peers at a single point in time and across time. For example, the technical manuals for Star Reading and Star Math (Renaissance Learning, 2011) report moderate to strong correlations with many state tests. Additionally, the reliability and concurrent validity of Star measures were found to be consistently moderate to strong as reviewed by the National Center for Response to Intervention (2010b).

In addition, CAT measures can examine change over short intervals consistent with progress monitoring. Probably most importantly, Computer-Adaptive Tests provide instructional direction in decision making. These measures not only tell educators how a student is doing, but have a distinct advantage over other measures, including CBM, in pointing educators toward next steps.

The remainder of this paper will examine and illustrate the use of one particular set of CAT measures, the Star assessments developed by Renaissance. The application of these measures within the RTI framework will be contrasted to the CBM measures in reading and mathematics.
**Screening with CAT**

Star Early Literacy (Star-EL), Star Reading (Star-R), and Star Math (Star-M) by Renaissance are a suite of CAT measures that meet the requirements of RTI. Star generates screening reports to assist data teams in determining which students appear to be on track and off track toward grade-level goals. Star’s Screening report can be run with standard RTI categories (benchmark, on watch, intervention, and urgent intervention) or a state’s AYP categories. Both represent important points in the distribution that are predictive of future success or failure.

The Star Screening Reports in Star-EL, Star-R, and Star-M provide educators with clear information about the standing of any single student relative to his/her peers. The reports also provide a listing of the specific students in the grade who fall into each category, making decisions from the data useful for grouping students during data team meetings. These data are the basis upon which teams in an RTI model identify students in need of intervention.

Star Screening Reports can also be used to examine changes over time in grade-level performance. For example, **Figure 2** shows the outcomes of a fall and winter Star Reading benchmark assessment of grade 5. In Scenario A, there is an increase of students moving into the benchmark area (green) and a reduction in those at the highest level of risk (red). Scenario B shows the opposite outcome with overall scores moving in the opposite direction of what would be desired. This type of report gives educators insight into the effectiveness of intervention and also the effectiveness of Tier 1 core instruction.

**Figure 2.** Screening Reports for Star Reading in fall and winter, comparing positive and negative scenarios.
Screening with CBM

Also for fall benchmarking, data obtained from a CBM early literacy measure—Nonsense Word Fluency (NWF)—is shown for 1st grade in Figure 3. The data are divided into three categories: average or above average = benchmark (green); below average = some risk (yellow); and well below average = at risk (red). These categories correspond to scores at or above the 25th percentile (benchmark), between the 10th and 24th percentile (below average), and those below the 10th percentile (much below average). Each CBM measure reports a rate-based metric, such as number of words read correct per minute for R-CBM. With the CBM model, different measures are administered depending on the grade.

CBM reports can also be used to show change over time. Figure 4 presents fall and winter screening outcomes with R-CBM for 4th grade. Two scenarios are displayed. In Scenario A, increases of students in the average or above average category (green) are evident, showing substantial gain in reading performance. In Scenario B, increases of students scoring in the some risk (yellow) and at risk (red) categories are evident. Teams use these data to indicate the effectiveness of Tier 1 core instruction over time.
Progress Monitoring with CBM

Once a team decides a student needs supplemental instructional support, appropriate interventions are selected. In order to determine if the instructional support is successfully impacting the student’s learning process, progress monitoring is conducted frequently while the student is receiving the tiered support. Both Star and CBM measures provide a similar process for progress monitoring.

An example of progress monitoring with a CBM, Figure 5 illustrates progress monitoring for a 1st grade student using Nonsense Word Fluency. Because the student was found to be in need of supplemental instructional support in the beginning of grade 1, the team placed the student into an intervention focused on improving knowledge of letter-sound combinations. Weekly progress monitoring of the student on NWF was conducted since NWF is the measure that reflects overall growth in phonics for students in grade 1. The goal set by the team was for the student to reach 63 correct sounds per minute by the end of the school year, a level that reflects achievement at the 50th percentile for students among the national sample of CBM users of the NWF measure. Over the course of 34 weeks, this was an expected rate of improvement of 0.85 (Rate of Improvement = (63-34)/34 weeks = 0.85 sounds per minute per week).

Examination of the report in Figure 5 showed that the student made little initial progress across the first 5 weeks of intervention. After consulting with the data decision team, the intervention was shifted to focus more on blending and decoding, indicated on the figure by a vertical red line. Following implementation of the intervention change, the student’s rate of progress was greater than expected, reaching a level of 1.1 sounds correct per minute per week. The team examined the data periodically and the successful response to the intervention was evident. Similar progress monitoring processes are used for R-CBM as well as CBM Math Computation and Math Concepts/Applications.
Progress Monitoring with CAT

An important aspect of using Star for progress monitoring is the nature of Computer-Adaptive Testing. Because the measure adjusts the difficulty of the items presented to students depending on the accuracy of their responses, the items answered correctly reflect a broad range of skills acquired by students.

For students in need of intervention, Star measures are repeated frequently (as often as weekly), producing data that reflect if students are responding to the intervention. When students fail to respond adequately, educators need guidance as to where the learning is "breaking down." Star measures produce data reports that are similar to those produced in RTI models using CBM, however, because Star is a Computer-Adaptive Test, the data go beyond telling educators how effective their intervention is and suggest next steps for instruction.

Before progress monitoring with Star, educators use the Goal Setting Wizard to determine expected rate of improvement. Figure 6 illustrates goal setting with Juanita Vargas, a first grade student with a scaled score of 503, which represents the 18th percentile and places her in the category of needing intervention (for more about scaled score, see page 17). The team chose the moderate rate of improvement suggested by Star’s Goal Setting Wizard. If Juanita grows at this rate, she will reach the 27th percentile by January.

Star’s suggested rates of improvement come from a growth model, which is a very large database of student growth patterns. With growth models, educators can make informed decisions about their expectations for growth. Star Reading’s growth model includes over 4.5 million students. The growth model for Star Math contains almost 1.1 million students, and Star Early Literacy’s model includes almost 200,000 students. Many state tests also use growth models.

Figure 6. The Goal Setting Wizard suggests a personalized rate of improvement for each student using a research-based growth model.
To monitor Juanita’s progress, Star-EL was administered weekly as seen in Figure 7. At first, Juanita did not show a growth rate equal to her goal. The team decided to change the intervention after 5 weeks.

The change in intervention resulted in substantial improvement in Juanita’s performance that actually surpassed the level of expected growth initially set for her. Based on her winter universal screening, the team will examine Juanita’s data and decide if she is still in need of the tiered intervention support. Star-R and Star-M also provide similar opportunities to conduct progress monitoring for students who are receiving supplemental tiered interventions.

Star monitors progress using scaled scores. Star scaled scores represent a set of skills students have a high probability of knowing. As a result, educators can run reports that suggest the skills students scoring at 470 versus 400 should have mastered, thus identifying potential targets for intervention. This is a fundamental element of all CAT measures: Student performance is placed on a developmental scale such that all students who score at the same level, regardless of grade, are essentially at the same point in the progression of skills resulting in that score.

Figure 7. Progress monitoring on Star-EL for a student in need of tiered support.
Understanding the Star Scaled Score

To fully understand the Star measures and their potential use in RTI, one must understand the Star scaled score. A scaled score is an indication of a student's placement on a test scale. When we weigh ourselves, we rely on a scale of “pounds” to differentiate our weight, regardless of our age. If we weigh an infant or a middle-age adult, we can easily communicate the relative differences between their weight because there are particular weights that define normal given a person’s age. Although we weigh infants and adults on the same scale (i.e., pounds), our expectation for where they fall on the scale is different depending on their ages.

Star Reading and Star Math use a scale of 1400 scaled score points. Star Early Literacy’s scale ranges from 300 to 900, spanning pre-K–3. Because the Star scale spans the entire school life of a child, the Star scaled score tells us where in the learning process the student falls. Although all students are placed on the same 0 to 1400 scale, our expectation for where a child should fall on the scale is related to their grade in school.

As illustrated in Figure 8, the expected Star Reading scaled score varies by grade and also represents a smooth and gradually increasing level of performance across grades. For example, students at the 50th percentile in spring of 2nd grade have a 334 scaled score. In 3rd grade, the scaled score for the 50th percentile in spring is 436. In 4th grade, it’s 515, and so forth.

When Star is used for universal screening, the scaled score tells the educator how much a student is behind relative to his/her peers. For example, looking at Figure 8, a 4th grade student who achieves a scaled score of 400 in spring is just below the 25th percentile compared to other fourth graders. Assuming that we want 4th graders to achieve at least at the 40th percentile in the spring, or a score of 470, educators can see there is a gap between where a student should be and where they currently are functioning, putting the student at some risk for academic failure.

<table>
<thead>
<tr>
<th>Grade</th>
<th>Percentile</th>
<th>Fall September</th>
<th>Winter January</th>
<th>Spring May</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Scaled Score</td>
<td>Est. ORF*</td>
<td>Scaled Score</td>
</tr>
<tr>
<td>1</td>
<td>10</td>
<td>59</td>
<td>5</td>
<td>70</td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>64</td>
<td>9</td>
<td>76</td>
</tr>
<tr>
<td></td>
<td>25</td>
<td>66</td>
<td>11</td>
<td>78</td>
</tr>
<tr>
<td></td>
<td>40</td>
<td>72</td>
<td>15</td>
<td>88</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>78</td>
<td>19</td>
<td>99</td>
</tr>
<tr>
<td>2</td>
<td>10</td>
<td>84</td>
<td>24</td>
<td>106</td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>100</td>
<td>30</td>
<td>161</td>
</tr>
<tr>
<td></td>
<td>25</td>
<td>110</td>
<td>32</td>
<td>181</td>
</tr>
<tr>
<td></td>
<td>40</td>
<td>166</td>
<td>43</td>
<td>232</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>197</td>
<td>51</td>
<td>263</td>
</tr>
<tr>
<td>3</td>
<td>10</td>
<td>184</td>
<td>49</td>
<td>222</td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>236</td>
<td>57</td>
<td>274</td>
</tr>
<tr>
<td></td>
<td>25</td>
<td>257</td>
<td>62</td>
<td>294</td>
</tr>
<tr>
<td></td>
<td>40</td>
<td>310</td>
<td>73</td>
<td>352</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>344</td>
<td>80</td>
<td>384</td>
</tr>
<tr>
<td>4</td>
<td>10</td>
<td>266</td>
<td>61</td>
<td>291</td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>321</td>
<td>73</td>
<td>351</td>
</tr>
<tr>
<td></td>
<td>25</td>
<td>344</td>
<td>79</td>
<td>372</td>
</tr>
<tr>
<td></td>
<td>40</td>
<td>402</td>
<td>94</td>
<td>441</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>445</td>
<td>103</td>
<td>475</td>
</tr>
</tbody>
</table>

Figure 8. Example of Star Reading scaled scores. Star Math and Star Early Literacy also span across grades.

* Est. ORF: Estimated Oral Reading Fluency is only reported for grades 1-4.
Instructional Planning with Star

The Star assessments offer information about skill development by signaling where students fall on the Core Progress learning progression linked to the assessment. A learning progression represents the prerequisite skills required to move through specific skills. Curriculum-Based Measures point educators to the need to conduct such diagnostic work, whereas Star provides the specifics of instructional planning as part of the routine assessment process.

Figure 9 shows the winter report from Star-EL of a hypothetical 1st grade student, Juanita Vargas. Juanita’s scaled score of 583 places her below benchmark, in the range identified as a late emergent reader. As a result, the RTI data team identified her as someone for whom supplemental intervention was needed. The report shows the nature of skill development embedded within each of the literacy domains assessed by the Star measure. For example, Juanita’s report shows that in phonemic awareness, blending word parts and phonemes would be the right targets for skill development. Within comprehension, reading and understanding words and complete sentences are focal points for intervention development. The report shows that she has generally good readiness skills and these skills need not be emphasized within the intervention.

Star also provides reports for instructional planning. The Instructional Planning reports in Star are based on the Core Progress learning progression and suggest the skills students are ready to learn next.

An examination of the Instructional Planning Report in Figure 10 shows the extent of instructional detail offered by the Star-M measure. In this case, the student, Brandon Bollig, has a score of 588, which is in the “on watch” area. The Instructional Planning Report identified specific skills within domains of 4th grade math on which instruction should be focused. For example, in Algebra, the Star assessment recommended an emphasis on determining the operation needed for a given situation as well as determining a multiplication or division sentence for a given situation. If Brandon’s learning is not accelerated, he will remain in the on watch area by the end of the year as predicted by Star.
Discussion

The purpose of this paper was to examine the premise that the CBM system equates to implementing RTI. As shown throughout the paper, this is a myth. The nature of the measurement system does not define the RTI model. RTI is about sound decision making and targeted instruction based on good data—this can be accomplished with several assessment systems including CBM and Computer-Adaptive Tests.

Both Star and CBM measures can work within RTI models to provide universal screening and progress monitoring. Certainly, there are advantages and disadvantages of each system. One advantage of Star is the link to instruction embedded in the assessments. As a result, there is more guidance about what could be done to improve student performance. Conversely, CBM requires educators to determine the instructional implications of a student’s performance. An advantage of CBM over Star is the length of the assessment. Whereas CBM measures usually take anywhere from 1 to 8 minutes depending on the domain being assessed, Star measures usually take between 10 and 25 minutes. This difference in time is not as large, however, when one adds the extra time needed with CBM to consider the nature of the instructional programs that are implied by the student performance.

Another difference between the measures is that CBM reading measures (Early Literacy and R-CBM) require individual administration by trained school personnel. CBM math measures can be administered to small groups of students. Star measures are computer administered. Although both measures involve educational personnel in the administration process, CBM measures tend to be more person-intensive than Star measures. Certainly, with Star measures, school personnel must be sure that students remain fully engaged with the computer during the assessment process. However, there is significantly less burden on school personnel in terms of administration time—especially for schoolwide screening three times a year.

A particularly important difference between CBM and Star measures is the nature of the measure. CBM in reading uses measures of fluency (correct per unit time) rather than accuracy (number correct) to reflect overall performance. In reading, a student’s overall performance is reflected in the rate at which he/she performs on the CBM. CBM math tends to use total correct (accuracy) but it can also be used as a fluency measure because one examines total accuracy per unit of time. In contrast, Star measures reflect a student’s accurate response to items representing various skills. Because Star is computer adaptive, each test adjusts based on the accuracy of the student’s response to questions representing multiple skills. As a result, the Star measures can give us substantial information about the nature of specific skills that students possess. The emphasis on skills provides opportunity for instructional planning that can be derived directly from the Star measures. This approach is very different than assessing individual skills according to the number of tasks completed correctly in a certain amount of time.
Conclusion

At the end of the day, users of Star measures can be confident that every aspect of RTI is easily met by the various reports generated through the Star assessments. Users of Star need to become comfortable and fluent with the Star scaled score, which on the surface, can seem a bit abstract compared to CBMs direct measures of student performance (i.e., the number of correct sounds, the number of words read per minute, the number of items correct on a math test). However, the Star scaled score represents where a student's skills fall across a longitudinal scale spanning the entire school spectrum from kindergarten through grade 12. Placement of students on a longitudinal scale is not part of the CBM measurement system.

Those considering Star measures should be confident that the data produced by Star assessments are accurate, reliable, and valuable for informing decisions that are a part of the RTI process. The Star measures provide a level of instructional planning information that exceeds what is produced by CBM alone, a level of support educators will certainly find useful within the RTI model. Star offers the added and important advantage of links directly to instructional targets compared to CBM.

The objective of effective RTI systems is to provide instructional interventions to students who are identified to potentially not meet grade-level expectations. Assessment tools are critical to RTI. Both Star and CBM systems measure student performance on the key components of universal screening and progress monitoring to reflect a student's response to intervention. Clearly, RTI does not equal CBM. Star measures offer an important and potentially valuable contribution to RTI.

Star measures offer an important and potentially valuable contribution to RTI.
References


Acknowledgements

Author
Edward S. Shapiro, Ph.D., is professor of school psychology and director of the Center for Promoting Research to Practice in the College of Education at Lehigh University. He is the 2006 winner of the American Psychological Association’s Division of School Psychology Senior Scientist Award. Professor Shapiro has authored 14 books and is best known for his work in curriculum-based assessment and interventions for academic skills problems. Among his many projects, Shapiro recently completed a federal project focused on the development of a multi-tiered, RTI model in two districts in Pennsylvania, and currently directs a U.S. Department of Education grant to train school psychologists as facilitators of RTI processes. He also collaborates with the Pennsylvania Department of Education in developing and facilitating the implementation of the state’s RTI methodology.

Reviewers
Matthew K. Burns, Ph.D., is a professor of educational psychology, coordinator of the School Psychology program, and co-director of the Minnesota Center for Reading Research at the University of Minnesota. Dr. Burns has published over 150 articles and book chapters in national publications, and has co-authored or co-edited several books. He is also the editor of School Psychology Review and past editor of Assessment for Effective Intervention. Specific areas in which Dr. Burns has conducted research include response to intervention, assessing the instructional level, academic interventions, and facilitating problem-solving teams.

Pat Quinn is known nationally as “The RTI Guy” and is the author of a widely circulated newsletter dedicated to helping teachers implement Response to Intervention (over 10,000 subscribers). His online training—“Response to Intervention Made Easy”—has been used by thousands of teachers around the country. Mr. Quinn’s latest book is titled Ultimate RTI.

Mike Vanderwood, Ph.D., is currently an associate professor of school psychology at the University of California–Riverside. He conducts research related to multi-tiered systems and English Language Learners. Professor Vanderwood has been involved in education reform for over 15 years and has helped implement multi-tiered systems in general and special education throughout the country. Most of his current research focuses on assessing and improving the quality of assessment and intervention tools used in a multi-tiered approach.

Jim Ysseldyke, Ph.D., is Emma Birkmaier Professor of Educational Leadership in the Department of Educational Psychology at the University of Minnesota. Professor Ysseldyke has been educating school psychologists and researchers for more than 35 years. He 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. 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. He chaired the task forces that produced the three Blueprints on the Future of Training and Practice in School Psychology, and he is former editor of Exceptional Children, the flagship journal of the Council for Exceptional Children. Ysseldyke has received awards for his research from the School Psychology Division of the American Psychological Association, the American Educational Research Association, and the Council for Exceptional Children. The University of Minnesota presented him a distinguished teaching award, and he received a distinguished alumni award from the University of Illinois.