
Technology-Based Assessment and Problem Analysis

Gerald Tindal and Julie Alonzo

This chapter discusses the features of a widely adopted technology-based assessment commonly used in schools implementing response to intervention (RTI). First, the chapter describes easyCBM[®] in the context of RTI. An important perspective is that a number of dimensions need to be addressed such as multiple references for decision-making, flexibility in measurement implementation, and documentation of technical adequacy. The second section addresses specific research conducted in the development of easyCBM[®] for reading and mathematics (organizing the discussion by grade bands: kindergarten to grade 2 and grade 3 through 8). Finally, this section ends by noting the strengths and utility of technology-based assessment to enhance decision-making. A number of reports are discussed that follow a sequence of teacher decision-making: (a) determining risk and grouping students, (b) diagnosing instructional needs of students and devising instructional plans, (c) evaluating these interventions using time-series progress monitoring, and finally (d) evaluating programs for both individuals and systems using grade-level and building-level changes, though a number of other variables could be considered (e.g., race–ethnicity, English language learners). The third and last sections of the chapter both reflect on critical dimensions of technology-based assessment and speculate on future directions in

this field. After articulating a conceptual model for easyCBM[®] using three interlocking critical features (measurement sufficiency, instructional adequacy, and data-driven decision-making), we venture into the need for professional development with effective reports to catalyze training enhancements.

Overview and Description of easyCBM[®]

Technology-based assessments have increasingly become essential tools for schools and districts implementing RTI. The best examples of technology-based assessment systems for use in RTI approaches share many features.

1. The process is efficient for staff to roster their students and activate the program with a minimal amount of work and yet allow flexible grouping of students, with the ability for teachers to regroup students throughout the year as individual needs change in relation to their peers.
2. Online platforms provide easy access to test themselves as well as the possibility of embedded trainings on test administration.
3. Well-designed databases facilitate immediate reporting of results, with historical records allowing districts to track the effect of different approaches over time. Such systems capture detailed information about the interventions programs and outcomes that provide critical information to guide lesson planning.

G. Tindal (✉) · J. Alonzo
University of Oregon, Eugene, USA
e-mail: gerald@uoregon.edu

4. The system is easy to use as the more complex computer programming is in the background and entirely managed using well-designed computer interface for all student and teacher interactions.
5. The system is designed for incremental enhancements to take advantage of advances in both measurement and technology.

This chapter focuses primarily on easyCBM[®], one such technology-based assessment system, designed as part of a federally funded National Center on Progress Monitoring. Although other technology-enhanced assessment systems share some of the same features, easyCBM[®] is used as the illustrative example because we are most familiar with its features and the principles that guided its design and fuel its continued enhancements.

Initially introduced in the fall of 2006, easyCBM[®] has continued to grow in popularity and use. At the time of this writing, over 325,000 educators, representing over 2.3 million students, have established accounts with the online learning system, with accounts present from every US state as well as a number of international locations. Over 17.4 million easyCBM[®] tests have been taken since the system was first made available. Many school districts have incorporated easyCBM[®] as an integral part of their RTI protocols. Student performances on the easyCBM[®] benchmark measures are used to identify students for additional intervention, and their scores on the easyCBM[®] progress-monitoring measures are used, in part, to evaluate the effectiveness of provided interventions and to modify instruction as needed. For academically struggling students, lack of progress on curriculum-based measurement (CBM) measures, when the students have been provided with appropriate intervention, implemented with integrity for a sufficient period of time, has served as the primary factor for determining eligibility to receive special education services.

CBMs Applied in the Context of RTI

CBM assessments include both individually administered measures (e.g., phoneme segmenting,

letter names, letter sounds, word reading fluency, and passage reading fluency) and group-administered measures (e.g., vocabulary, comprehension, mathematics). Test administration procedures are typically described in a teachers' manual. On the easyCBM[®] system, the utility of the teachers' manual, available as a downloadable portable document format (PDF) directly from the easyCBM[®] site, is bolstered by the addition of an online training link, with videos on standardized test administration and scoring and proficiency checks of mastery to be completed before a person administers the tests. For greatest utility, assessment systems designed for use in RTI contexts must include both screening and progress-monitoring measures.

As is typical of such assessment systems, the easyCBM[®] system includes both universal screening measures (for fall, winter, and spring administrations) and multiple alternate forms used in evaluating instruction and progress. Specific measures included on the screening assessments vary by grade level, with the measure type selected based on empirical findings of which tests provide the most robust screening of content and skill at each grade level.

An array of progress-monitoring assessments should be available at each grade. These measures provide in-depth information about individual students' particular strengths, weaknesses, and needs for tier 2 supplemental instruction and tier 3 intensive intervention. Progress-monitoring forms typically include not only the specific measure types included on the universal screening assessment but also go beyond these measures to provide information about a wider swath of skills. These diagnostic measures should not only provide relevant information about specific skill areas with which students are struggling but they should also be optimized to be sensitive to growth, enabling teachers to evaluate the impact of their instruction and modify it when warranted. In the most useful technology-based systems, intervention data are logged directly into the system, which plots intervention lines on the individual student graphs, providing individualized histories of student response to instruction or tiered supports.

When systems truly capitalize on the strengths of technology-based assessments, the group-administered measures from both universal screening and progress-monitoring assessments can be administered via desktop computers, laptop computers, or tablet devices such as iPads. Individually administered measures can be administered using paper and pencil, with scores later entered online, or they can be administered directly from tablet devices, to streamline data collection.

Technical Adequacy Considerations

In keeping with the original precepts of CBM, the measures need to be sensitive to the instructional needs of students with disabilities as well as to students from the general education population. Because of the need for alternate forms, measures need to be appropriately scaled, a consideration that can be addressed during measurement development with the use of item response theory (IRT). Finally, traditional reliability and validity requirements need to be met, as promulgated by the 2014 *Standards for Educational and Psychological Testing* (American Educational Research Association, American Psychological Association, and National Council on Measurement in Education 2014).

For example, all easyCBM[®] measures were developed using the principles of universal design for assessment (Thompson et al. 2005) and followed the guidelines for test development as described in *The Standards for Educational and Psychological Testing* (American Educational Research Association, American Psychological Association, and National Council on Measurement in Education 2014) with particular attention to accessibility and freedom from bias. During item development, teams of experienced grade-level educators were hired to draft the measures, with thorough review by a team of trained assessment researchers at the University of Oregon. Following content and bias/sensitivity review, items were piloted with grade-level students (with exact sample size varying by measure and grade). IRT modeling was used to analyze pilot data, with item information (measure, standard

error, discrimination, mean square outfit) used to create multiple alternate forms for screening and progress-monitoring measures.

Screening assessments are typically optimized for the purpose of identifying students who need additional supports to meet grade-level content and performance standards in the areas of reading and mathematics. Progress-monitoring assessments, on the other hand, may be designed to provide more detailed diagnostic information related to specific skill deficits and developed to ensure comparability of alternate forms and sensitivity to growth for students receiving targeted instruction in specific skill areas.

Following initial form creation, additional studies are typically conducted to evaluate and document the reliability (test–retest, alternate form), internal consistency (both within and between measure types), sensitivity and specificity, generalizability of the measures, and to provide evidence of their validity in making screening- and progress-monitoring decisions. Results of these studies should be detailed in technical reports or other published materials readily available from the CBM developers. For instance, over 100 technical reports documenting easyCBM[®] measurement development and technical adequacy studies are available on the website of behavioral research and teaching (BRT) at the University of Oregon (brtprojects.org). The need to consider the technical characteristics of the measures is emphasized because it is believed that instructional planning and evaluation is not a low-stakes decision that can be made on the basis of informal measures. Rather, it is a high-stakes decision that needs to use measures that have all the characteristics demanded by the test standards.

Technology-Enhanced Reading Measures In kindergarten, first, and second grade, early literacy measures focus on letter names, letter sounds, phoneme segmenting, and word-reading fluency (Alonzo and Tindal 2007); in grades 2 to 8, vocabulary and direct measures of comprehension measures are available (Alonzo et al. 2012a-g; Irvin et al. 2011, 2012a, b; Lai et al. 2010; 2012a, b; Park et al. 2012a, b). Although

other approaches are possible, use of IRT enables creation of item pool with known item difficulty, facilitating the creation of comparable alternate forms. For instance, use of IRT during measurement development resulted in the easyCBM[®] early literacy measures alternate form reliability ranging from 0.86–0.91 for the Phoneme Segmenting measures, 0.82–0.89 for the Letter Names measures, 0.76–0.88 for the Letter Sounds measures, and 0.95–0.96 for the Word-Reading fluency measures in grades K-2.

In the late elementary and middle school grades, a variety of reading measures have been developed and investigated, including word and passage reading fluency, vocabulary, and comprehension. Some CBM systems use maze tasks for measuring comprehension, while others use passages followed by selected response items. As with the early literacy measures described earlier, use of IRT during measurement development provides insights into item characteristics instrumental in creating assessments that are sensitive for use as screeners as well as for monitoring the progress made over time. With a view toward ensuring the utility of the assessment for students at risk, developers may focus on building a test that begins with a sufficient number of easily accessible items to establish an accurate base of knowledge or skill, yet also includes items that cross the range of difficulty, providing the means by which to sensitively measure improvements over time for students with a range of skill/knowledge. Measures that contain an insufficient number of “easy” items, or alternatively that contain an insufficient number of “moderately difficult” and “difficult” items may not be useful for screening purposes, and also fail to serve the dual purpose of providing information about progress over time.

Technology-Enhanced Mathematics Measures Although early work in mathematics CBM focused on fluency-based short probes, more recent developments reflect the increasing depth of knowledge and skill expected of students, with a greater emphasis on items aligned to state and content standards, such as the Common Core State Standards (CCSS) in Mathematics and the

National Council of Teachers of Mathematics’ (NCTM) Focal Point Standards.

The easyCBM[®] mathematics progress-monitoring measures were initially developed to be sensitive for students with persistent learning problems (Alonzo et al. 2006; Anderson et al. 2010, 2011), using the NCTM Focal Point Standards to establish the content and performance expectations used (Alonzo et al. 2010; Nese et al. 2010a, b). At the time this chapter was written, the easyCBM[®] mathematics measures were being revised with (a) alignment to the CCSS for the existing measures and forms and (b) development of new items and forms written in alignment with the CCSS. As in the reading measures, the same Rasch scaling was used to create equivalent alternate forms in grades K-z (Alonzo et al. 2009b, c, d). A series of technical reports were written describing the process of developing items, articulating the blueprint and scaling process as well as the development of forms: grade 3 (Alonzo et al. 2009e), grade 4 (Alonzo et al. 2009f), grade 5 (Lai et al. 2009c), grade 6 (Lai et al. 2009b), grade 7 (Lai et al. 2009b), and grade 8 (Lai et al. 2009d). Both students and items were placed on the same scale using a 1-parameter Rasch model; the results showed the average for items was lower than the average for students and alternate test forms within each grade had a mean difficulty (IRT *Measure*) within 0.20 Rasch units of one another (Anderson et al. 2011). Internal consistency ranging from 0.78–0.91 was documented once the alternate forms were completed (Anderson et al. 2009, 2012a, b). Validity also has been established (Anderson et al. 2011a, b, c, d).

Additional Technical Adequacy Considerations

School districts increasingly use technology-enhanced assessments to identify students who are at risk of not passing their statewide assessments. For this application, studies establishing the utility of the assessments for that purpose are crucial. Normative performance benchmarks can provide insights into how a student’s performance compares to grade-level

peers (Tindal et al. 2009a, b). A variety of criterion-related validity studies documenting the sensitivity and specificity of CBM assessments for predicting performance on statewide large-scale assessments have been published. Such studies typically include correlational as well as multiple regression analyses (Anderson et al. 2010a, b). In addition, it is becoming more common for assessment developers to report on the diagnostic efficiency of such measures. For example, the diagnostic efficiency of the easyCBM[®] mathematics measures was determined using a receiver operating characteristic (ROC) analysis; ROC analyses were cross validated with unique samples. In a ROC analysis, classification of students (using state test proficiency status) is used to determine sensitivity (true positive proportion) and specificity (true negative proportion) of the screening (benchmark) measures. A ROC curve plots the screening test's false-positive rate on the horizontal axis and sensitivity on the vertical axis for $k-1$ scores where k =the total number of unique scores occurring on the screening measure. Thus, each point on the curve depicts the sensitivity and false-positive rate for every score in the range of scores on the test measure (<http://www.ajronline.org/content/184/2/364.full>) reflecting the reality that gains in sensitivity always come with increased false-positive error rates. Area under the curve (AUC) curve are used as a measure of predictive power. Results indicate that the easyCBM[®] math measures function well, with AUC results ranging from 0.85 to 1.00 across grades K-8 (Anderson et al. 2011a, b).

In summary, technology-enhanced assessment systems such as easyCBM[®] have benefited from years of research funded by the Institute of Education Sciences (IES) and, earlier, the Office of Special Education Programs. Technical reports are available on the BRT website: <http://brtprojects.org>. This ongoing and systematic program of research continues to inform refinement of such assessment systems. In-depth longitudinal study of the measures in use, along with regular feedback solicited from practitioners using the measures in their local RTI contexts ensures continued development and adaptation to meet the needs of educators.

Strengths of Technology-Based Assessments

Technology-based assessments offer many benefits to users including increasing the security and efficiency of data collection efforts, streamlined processes for sharing results, and savings of time and money by reducing the need for printing and shipping testing materials. In addition to these more obvious benefits, technology-based RTI systems can offer some more subtle insights, such as facilitating deeper understanding of student performance data, fostering a team approach to meeting student needs, and harnessing the power of databases to assist in identifying patterns in the data that might not be so readily observable from visual inspection of raw scores.

Similar to other technology-based assessment systems, the easyCBM[®] system uses secure uniform resource locators (URLs), with administrator control of access to data. District-level users are able to access performance information from all users in a school district; building-level users see performance data from all students assigned to their school, and classroom-level users can access data only for students specifically associated with them in the database. Fluid-grouping features enable multiple users to access the data and form groups for analysis/interpretation without disturbing other users' organizational structure. These features enable a Title 1 coordinator or a special education professional, for example, to access student performance records for students in their caseloads, even when others (e.g., classroom teachers, instructional assistants) administer the tests. Each user is assigned a unique user name and password (which they can update to maintain security), and data are encrypted such that they cannot be viewed except when using the reports on easyCBM[®] or exported in comma separated values (csv) format in preparation for upload to a district or state student information system.

Reports on the technology-enhanced assessment systems are designed to provide useful information to guide decision-making to meet different users' needs. At the district level, reports can facilitate discussions related to resource allo-

cation and program evaluation, enabling district administrators to document the impact on student learning of district initiatives and identify areas of greater/lesser need, and specific skill areas for which professional development or additional supports might be needed to enable students to meet rigorous content and performance standards. At the school level, reports enable users to identify specific broad constructs (vocabulary, fluency, comprehension, mathematics) in which students are struggling, thus facilitating decisions related to programmatic and curricular supports. Classroom-level reports provide insights into the specific skills students have mastered or with which they are struggling (e.g., specific letter sounds, particular objectives within math, or inferential rather than literal comprehension), fostering informed lesson planning based on student needs. Individual reports enable teaching teams to monitor the effectiveness of specific interventions for individual students and groups of students and provide an accessible and efficient way to communicate with parents.

One of the greatest benefits of using a technology-based assessment system is the increased functionality it provides in terms of tracking student responses and providing useful feedback to teachers and students themselves. Technology-enhanced assessment systems optimize group-administered measures for online testing, with student responses captured in real time as students complete the tests, enabling instant score reporting upon completion of the assessment. Carefully thought-out systems can enable teachers to track student completion of assessment items as they are working their way through the tests, and can provide a variety of reports and graphs to help teachers interpret student performance without having to engage in additional data entry or recording of student data. Programs can include detailed reports for students, helping them track their own progress in specific skill areas and providing instant feedback on their performance, enabling them to identify areas of confusion that would benefit from additional attention in their studies.

Well-designed computer-administered assessments can also streamline the testing experience for students, with accommodations built into the infrastructure such that they can be delivered with a minimum of demand on school resources. For instance, on assessments that do not target reading, audio files can be embedded within the website, enabling students to access a read-aloud accommodation without needing to be singled out for individual attention or one-on-one administration. All of the easyCBM[®] mathematics items that include words, for instance, include a read-aloud option. And, as of January 2014, all easyCBM[®] mathematics items include an optional Spanish language accommodation enabling district administrators to provide Spanish language accommodations (both written text and read aloud) for every mathematics item K-8. Again, forethought about programming is essential. In the case of the Spanish language accommodation, for instance, it is important that administrators be able to deactivate this option in states where Spanish language accommodations are not allowed, and reports must be programmed to record whether the test was taken in an English-only or a language-accommodated condition.

In addition, with intentional planning and design, features such as careful use of blank space to enhance accessibility and attention to the fine motor skills required of students as they are interacting with an online assessment can be integrated throughout. Such design considerations play an important role in reducing construct-irrelevant variance that might be introduced for students who lack dexterity with the computer or who have visual challenges, reducing the likelihood that such student characteristics will cause scores to be artificially deflated. The technology used in some online assessment systems enables the measures to be made available as downloadable PDFs. The ability to download the measures for students who have difficulty accessing them in their online format facilitates local accommodations to meet IEP requirements and can also accommodate settings with limited computer access. Many online tests can be taken in a single sitting or administered in shorter segments to ac-

commodate student needs. When this technology is included, student responses can be retained in the system, with the computer automatically returning the student to the test at the point at which he/she stopped, if multiple testing sessions are used.

Ways Technology-Based Assessment Enhances Decision-Making

Computer-administered assessments offer additional benefits in the form of instantaneous scoring and report generation. When attention is paid to the need for quick access to testing results, systems can be programmed to provide reports as soon as students finish their assessments. Computer-administered assessments can also streamline the process of providing opportunities for students to make up tests they have missed due to absence, enabling teachers to administer the make-up assessments in a group or individual setting, based on which provides the most benefit in a given situation.

Thus, well-designed technology-based assessments can provide criterion-referenced information (as with the item-level analysis shown in Fig. 2), individually referenced information (as in the individual progress monitoring graph shown in Fig. 4), as well as norm-referenced information (as shown in Figs. 6–8). Benchmark screening reports, such as the one displayed in Fig. 6, help teachers quickly identify which skills individual students are struggling with as well as to identify patterns in their classes as a whole. Usability features such as making the data sortable from high to low or low to high by clicking on the column headers enhance the utility of such reports. The next section describes a sequence of decisions that technology-enhanced assessments facilitate.

Risk Analysis and Group Assignment One of the initial decisions that need to be made at the beginning of the school year is to determine who is at risk for learning problems. For this decision, benchmark measures are administered to all students during fixed periods of time (fall, win-

ter, spring). These common assessments allow teachers to compare students to each other and determine where they fall in a performance distribution. Students who are performing below district-set percentile rank (PR) cut points can be identified. When a comparatively low level of performance appears across different measures, it likely represents an overall need for extra support, classified on the easyCBM[®] system as a *risk rating*. In Fig. 1, students are grouped on the easyCBM[®] fall measures into four groups: (a) below the 20th PR in all three areas measured, displayed as red; (b) below the 20th PR in a couple of measured areas, displayed as yellow; and (c) not below the 20th PR in any area measured, displayed with no color; and (d) exceptionally high performance (e.g., 90th PR) in any measured area, displayed in green. In the end, a risk classification is given as *high* (those marked red), *some* (inconsistently red and yellow), or *low* (few to no areas of risk). See Fig. 1 for a sample display of students in a teacher's classroom.

Group reports help teachers make decisions about how to organize their students for instruction. Figure 2, for instance, displays the performance of a whole class on the fall reading comprehension test. The bar graph highlights the difference in proficiency in reading comprehension of students in the particular class, enabling teachers to identify the three students who may benefit from a more intensive intervention (or may need a fluency-building intervention to assist in comprehension) as well as the seven students whose performance indicates a need for targeted comprehension instruction.

Design of Instruction with Attention to Detail

Widespread access to educational technology has expanded the universe of possibilities for effective assessment systems in support of RTI. Some features of technology, in particular, are worth highlighting as they relate to assessment considerations. Computers and tablet devices such as iPads provide opportunities to capture item-level data that can provide additional instructionally relevant information, over and above raw or scaled scores. For instance, it is possible to use item-level responses, captured

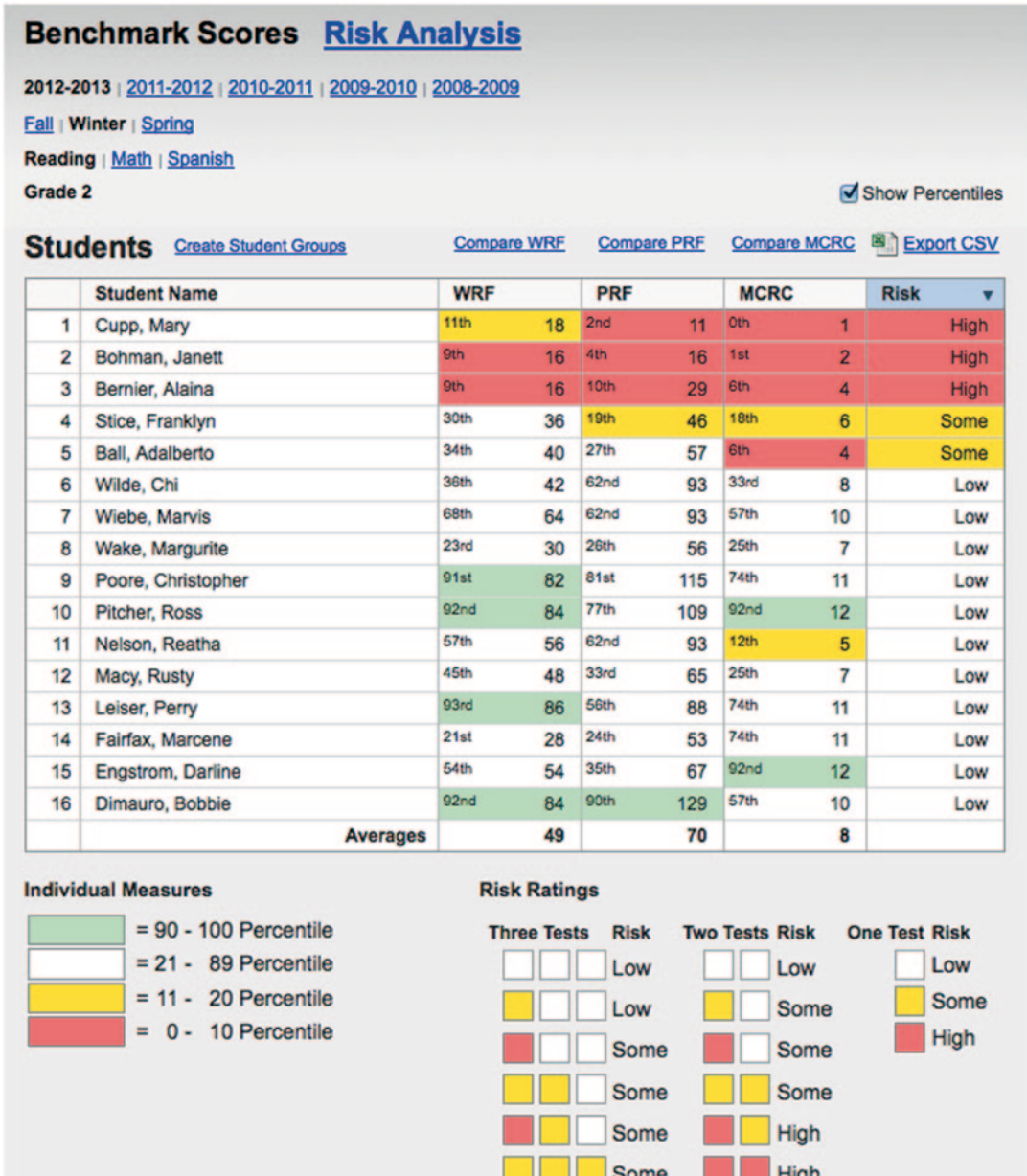


Fig. 1 Group benchmark report

via online test administration, to provide reports that document specific areas of weakness or strength for students in a given classroom. Figure 3, for example, shows a report from one of the easyCBM® comprehension measures. This report gives the teacher insight into the need to focus more on inferential and evaluative comprehension than on literal comprehension when

planning lessons, and also highlights the specific error patterns shown by different students in the class. Computer programming within the system enables a teacher to identify patterns in individual student responses: When a teacher “hovers” over a student’s name in the list, the name is highlighted, making it easier for teachers to find

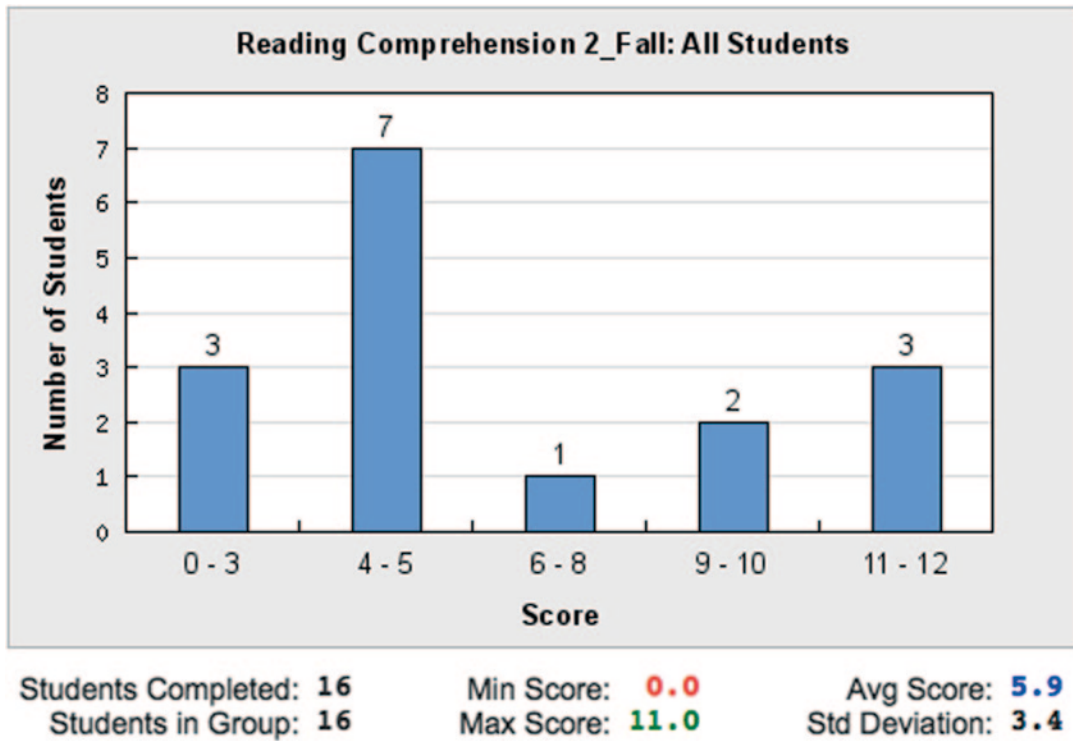


Fig. 2 Group report on a fall benchmark screener assessment

patterns in student responses they might not otherwise notice.

Intervention Evaluation A key feature of RTI is the focus on individual students' response to interventions. Technology facilitates the gathering and interpretation of such information. Figure 4 shows a screen shot from an easyCBM® individual student progress report.

Student scores are plotted over time, enabling the teacher to evaluate growth across the school year. Colored lines on the graph provide normative referents, enabling quick evaluation of the degree to which the student's performance is at grade level (meets the 50th percentile line plotted on the graph) or falls below expectations (to the degree that a student's performance falls further below the 50th percentile, his/her level of need can be interpreted to be greater). These individual reports also provide a place to record the interventions a student has received, with the database providing a convenient and cost-effective historical archive of these data, so subsequent

years' teachers have ready access to the instructional approaches that have proven effective—or ineffective—for a given student. The utility of this historical record that combines details about the interventions provided and a record of student performance over time cannot be overemphasized. Without the log of interventions (See Fig. 5), assessment data are at best, difficult to interpret and at worst, relatively meaningless.

Program Evaluation for Individuals and Systems Figure 6 documents an individual student's assessment performance history against the backdrop of the district's performance on the same assessments. At a glance, one can see that the example student Adalberto's oral reading fluency in the fall of second grade (as measured on both the Word and Passage Reading Fluency measures) was near grade-level expectations, with scores that placed him at the 45th and 59th percentile, respectively, when compared to national norms, while his performance on the comprehension measure in the fall placed him near the cutoff

Item Analysis

| Easiest to Hardest Items | | | | |
|--------------------------|-------------|------------------|------------|--|
| Item | Type | Students Correct | Percentage | Student Names, Incorrect |
| 6 | Inferential | 4 of 4 | 100% | |
| 5 | Literal | 4 of 4 | 100% | |
| 1 | Literal | 4 of 4 | 100% | |
| 12 | Inferential | 3 of 4 | 75% | Christia T |
| 19 | Literal | 3 of 4 | 75% | Denver S |
| 14 | Literal | 3 of 4 | 75% | Denyse F |
| 16 | Evaluative | 3 of 4 | 75% | Christia T |
| 15 | Inferential | 3 of 4 | 75% | Denyse F |
| 8 | Literal | 3 of 4 | 75% | Denyse F |
| 3 | Inferential | 3 of 4 | 75% | Christia T |
| 2 | Literal | 3 of 4 | 75% | Christia T |
| 17 | Inferential | 3 of 4 | 75% | Christia T |
| 4 | Inferential | 2 of 4 | 50% | Gail P, Christia T |
| 7 | Evaluative | 2 of 4 | 50% | Denyse F, Christia T |
| 20 | Evaluative | 2 of 4 | 50% | Denyse F, Denver S |
| 18 | Evaluative | 1 of 4 | 25% | Gail P, Denver S, Christia T |
| 10 | Evaluative | 1 of 4 | 25% | Gail P, Denver S, Christia T |
| 9 | Inferential | 1 of 4 | 25% | Denyse F, Gail P, Christia T |
| 11 | Literal | 0 of 4 | 0% | Denyse F, Gail P, Denver S, Christia T |
| 13 | Evaluative | 0 of 4 | 0% | Denyse F, Gail P, Denver S, Christia T |

Fig. 3 Item analysis report to target instruction

point for being identified as at risk in his district, with a 23rd PR. By the winter benchmark assessment, although Adalberto’s raw scores continued to improve, his standing in relation to his peers dropped, placing him between the 27th and 34th percentile in oral reading fluency and at the 6th percentile in reading comprehension. This report can help prompt discussion and problem-solving on the part of Adalberto’s teachers, as they work together to put together an instructional plan that will help him reverse this downward trend in performance.

Technology-based assessments constructed on a foundation of well-designed databases facilitate district-level decision-making as well. Figure 7 depicts a school comparison evaluation report that provides administrators with a wealth of infor-

mation in a single screen. At a glance, they can identify the schools in their district with the most intense need for academic support structures, and by toggling back and forth between the “Counts” and the “Percentages” options, they can quickly gather data that might be relevant to share with stakeholder groups, including school boards and, potentially, outside agencies that might offer funding opportunities to support innovations. The links on the page enable users to focus on specific grades or all grades concurrently for content areas. Users can click on school name to dig deeper into patterns they want to explore based on the initial review of the data. Within the Building Level report, district administrators can review the progress students are making over the year, by class.

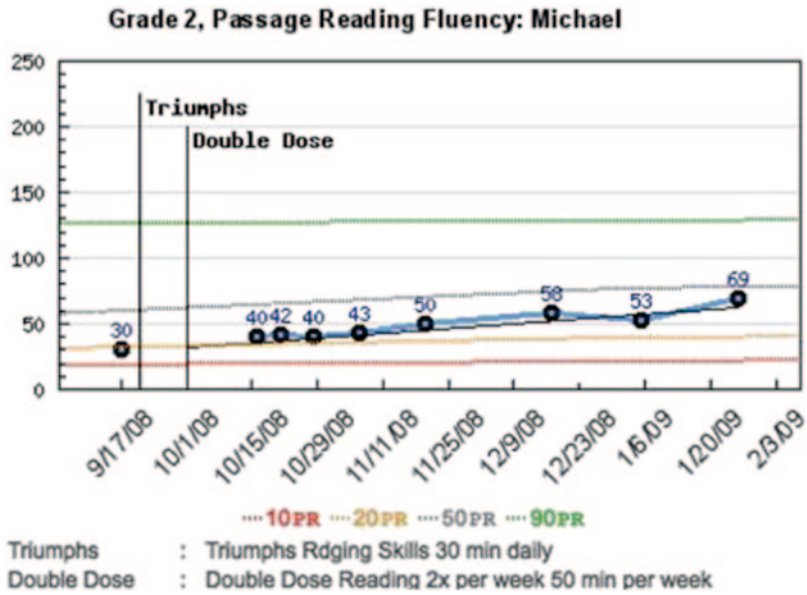


Fig. 4 Individual student progress report to evaluate instruction

easyCBM
Logged in: tealy (logout)
4/3/2013 - 2:41:35 pm

Home
Students
Measures
Reports
Account

Reports and Analysis

To view a **Group** report, click on the name of the group, and all of their active CBMs will appear below. Select a CBM name to see a summary and list of student scores. Then click "View" to see any student's actual submission. Select the **Individuals** subsection for easy one-click access to system wide data by student. **Interventions** are now accessible under the Individuals subsection, on the right-hand side of the table.

Benchmarks
Groups
Individuals

Students > Interventions for Mary Cupp

| Date | Subjects | Label | Description | |
|-----------|----------|---------|---|---|
| 9/15/2008 | Reading | Tier 2 | 30 minutes of HM differentiated materials in a small group. | Edit Delete |
| 10/6/2008 | Reading | BEST | 60 minutes 4 times a week of small group instruction using Soar to Success. | Edit Delete |
| 12/1/2008 | Reading | Phonics | 15 minutes a day of phonics instruction in a small group of 5 kids using Phonics for Readers. | Edit Delete |

New Intervention

Fig. 5 Detailed log to label and describe instruction for each student

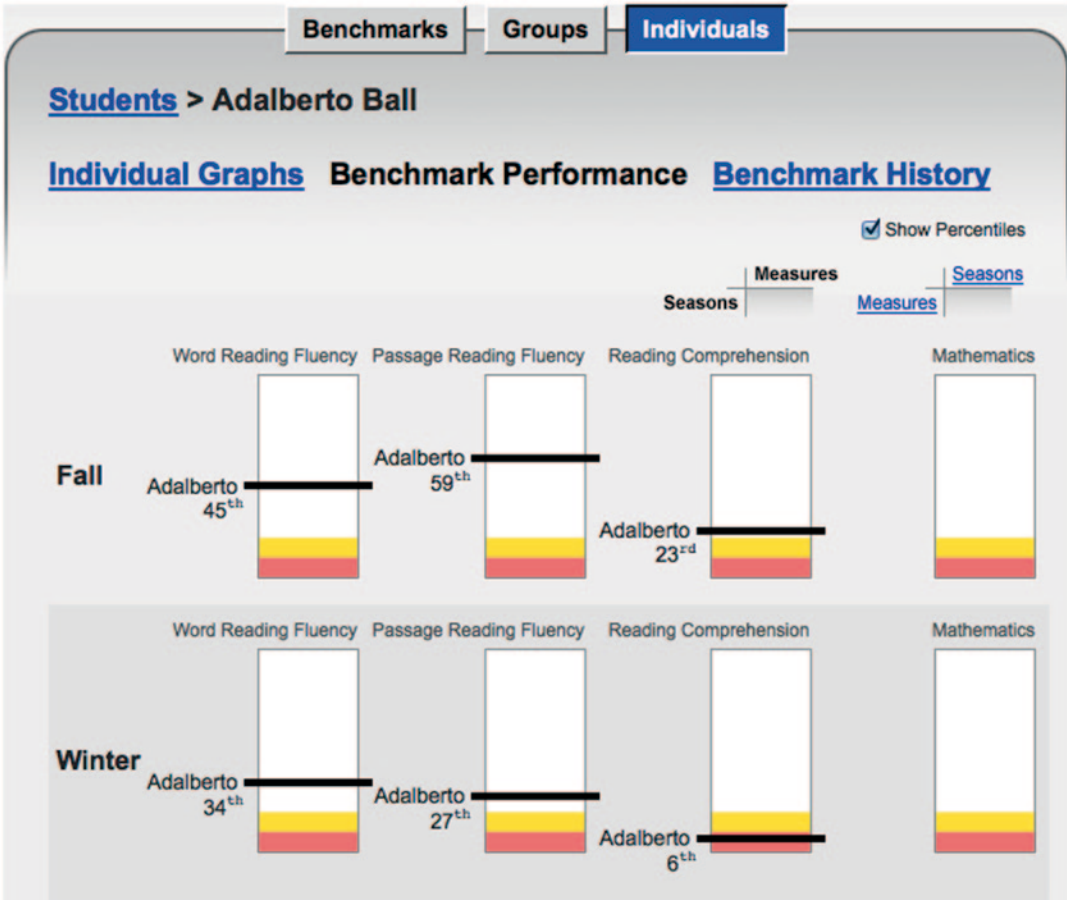


Fig. 6 Individual benchmark history report

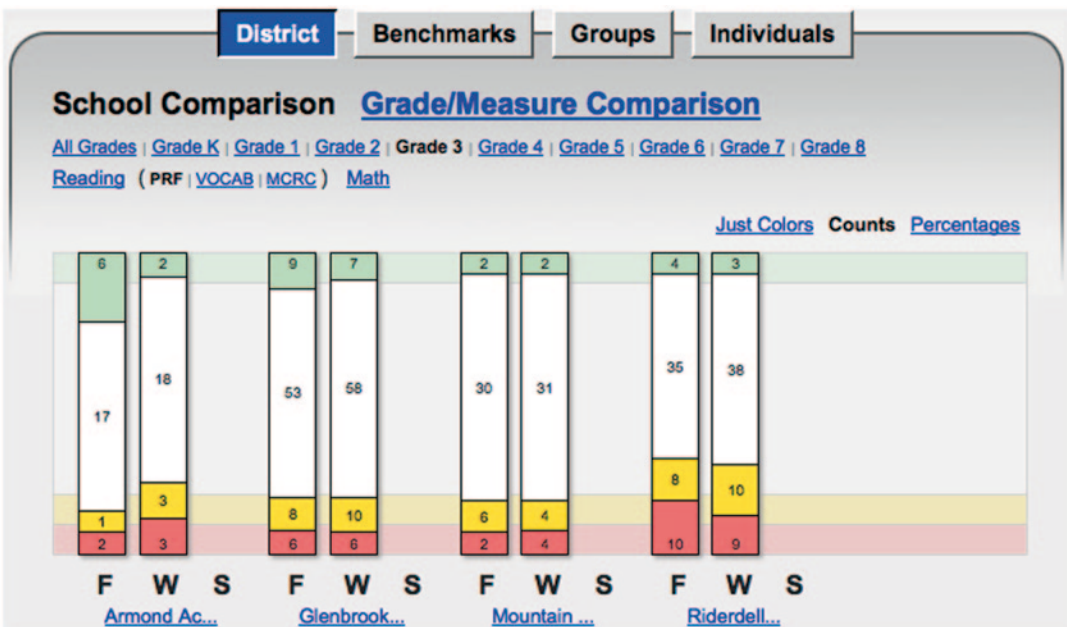


Fig. 7 Systems evaluation report with school outcomes

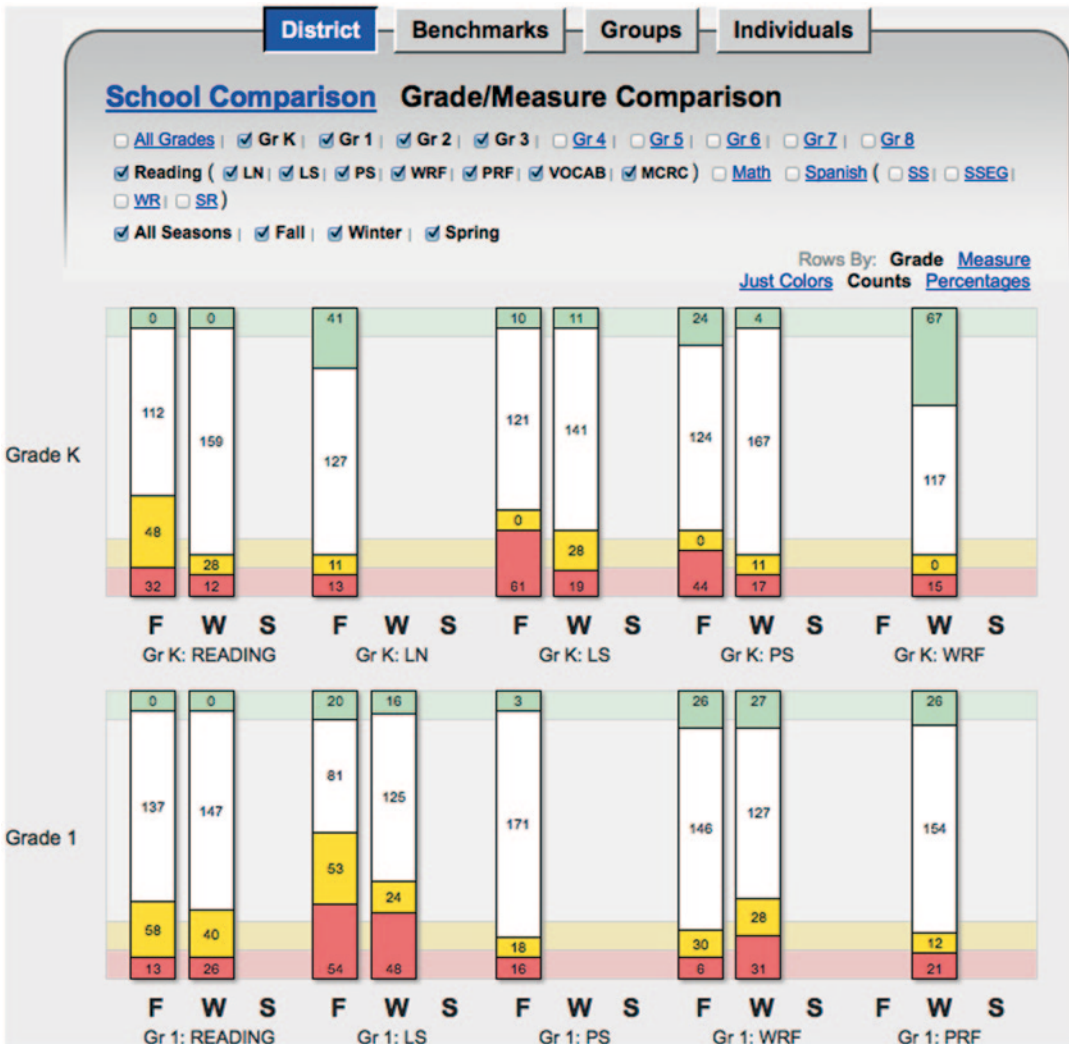


Fig. 8 Systems evaluation with grade/measure outcomes

Figure 8 displays the grade/measure comparison report available for district-level users. Using this report, administrators can look for patterns in their overall district performance, organized by grade level and season, with the ability to select measures. Color-coding, based on the district’s preset risk ratings, facilitates interpretation of the reports.

In summary, well-designed technology-based assessments can provide norm-referenced information like the benchmark screening reports, as was displayed in Figs. 1 and 2, helping teachers quickly identify who is at risk, facilitating grouping of students to maximize instructional oppor-

tunities. Once these macro outcomes are accomplished, further reports can be generated to target specific skills with which individual students are struggling as well as to identify patterns in their class as a whole, providing criterion-referenced information (shown in Fig. 3) that can be used to document specific instructional programs (Fig. 5). At this point, formative assessment can be used to evaluate instructional programs using individually referenced information (as in the individual progress monitoring graph shown in Fig. 4). After a period of time (perhaps within each seasonal assessment), programs can be

evaluated for individuals using norm-referenced information (as shown in Figs. 6–8).

Reflections and New Directions

The essential features of progress-monitoring assessments, often referred to as CBMs, have changed very little since the 1970s. Such measures are still expected to sample from a year's worth of curriculum. They are meant to provide teachers with meaningful information about the progress students are making in mastering that material. In addition, to enhance their utility, progress-monitoring measures are intended to be easy to administer, score, and interpret. See Tindal (2013) for a historical summary of CBM.

However, whereas four decades ago, researchers deemed CBMs as not requiring any particular expertise to develop, the increasing stakes associated with assessment results as well as advances in psychometrics have significantly altered this perspective: It is now recognized that the creation of reliable and valid progress-monitoring measures requires specialized knowledge beyond what most public school teachers possess. This realization spurred the creation of “next-generation” CBMs, measures created using rigorous alignment with standards and stringent statistical modeling (Alonzo et al. 2006). In all of our measurement development, the authors have worked with expert teachers throughout the country to develop reading and mathematics benchmark and progress measures.

Probably the most important advancement with easyCBM[®] and other modern technology-enhanced assessments is the use of IRT (Embretson and Reise 2000) during test construction, which sets these measures apart from more traditionally designed formative assessments (aimweb and Dynamic Indicators of Basic Early Literacy Skills; DIBELS). With IRT, test developers are able to calculate measure difficulty for every item and then develop forms that are equivalent. In using IRT, test developers also have been able to place items and students on the same scale to also ensure sensitivity to progress. Such mea-

asures stand out for their sensitivity in monitoring growth, the stability of alternate forms, and the provision of measures suitable to assess the full range of student skill in critical content areas from kindergarten through eighth grade.

Another significant difference between easyCBM[®] and most other progress measures has been the focus on alignment with standards. By design, items have been built in both reading and mathematics that are aligned with standards. When easyCBM[®] was initially developed, the National Reading Panel's report was used to develop measures in reading and the NCTM Focal Point Standards for measures in mathematics. More recently, however, new items have been developed that are aligned with the CCSS in both content areas. Formal studies of the alignment between these national standards and the assessments document the measures' appropriateness for use in standards-based school systems.

Conceptual Model for Impacting Learning with Technology-Enhanced Assessments Although scaling and alignment form two recent innovations into progress monitoring with easyCBM[®], the most significant innovation is the movement beyond a measurement model to an instructional model for which progress monitoring is meant to be used in evaluation (Tindal 2013). The conceptual model for easyCBM[®] is based on three warrants designed to optimize RTI:

1. Assumption 1 (measurement sufficiency): Students are appropriately placed in long-range goal material to ensure the measures are sensitive to change: What is the type of measure, grade level of measure, and the time interval (density of measures) used during the year?
2. Assumption 2 (instructional adequacy): Instruction is detailed and explicit, allowing a team of teachers to coordinate various elements such as providing an instructional tier (1–3), allocating sufficient time to teach, grouping students appropriately, deciding on the instructional emphasis (alphabetic principles and phonemic awareness, decoding, fluency, vocabulary, and comprehension), using specific curriculum (core and supplemental)

materials, and determining what instructional strategies to use.

3. Assumption 3 (decision-making): Interventions need to be introduced for low-performing students when data warrant change. Are interventions provided at the right time and in accord with specific data features (e.g., level, variability, and slope)?

The theory of change in using technology-enhanced assessments to drive improvement in learning is best reflected as the interlocking union of the three components in a chain: measurement sufficiency, instructional adequacy, and decision-making. It is not each link itself that is critical, but the intersection of the link with subsequent links. As teachers collect data (from benchmark to decisions of risk and monitoring of progress), the data used to inform them need to be sufficient, directed toward instruction, and adjusted as needed (flexible and prescriptive). Furthermore, this information needs to be collected into a single database for teachers to monitor their application of RTI as well as policy-makers and administrators to use the information in making systems-level decisions.

The conceptual model is driven by accessibility as a key ingredient to change: If information is not easily accessible and tractable, then it is unlikely to result in use. The theory is also driven by a holistic approach to change: Changing individual components as separate events is unlikely to change systems. Rather, the whole needs to be reflected in the parts that in turn need to connect teachers and administrators. Finally, research in practice is needed. By using a developmental process, the influence of all components can be modeled individually and integrally to establish optimal influences through a structural model (Fig. 9).

The combined effects from all three components (proximal variables) are critical as well as the relation between them and the outcome (distal variables), which is within-year growth (on benchmark measures), to document change relative to peers. It is not enough to have only one of the proximal variables—the right measures, targeted interventions based on best evidence, and decisions tied to their effect on students. All three are needed. However, they need to work synergistically. And even then, changing these three components is not enough either.

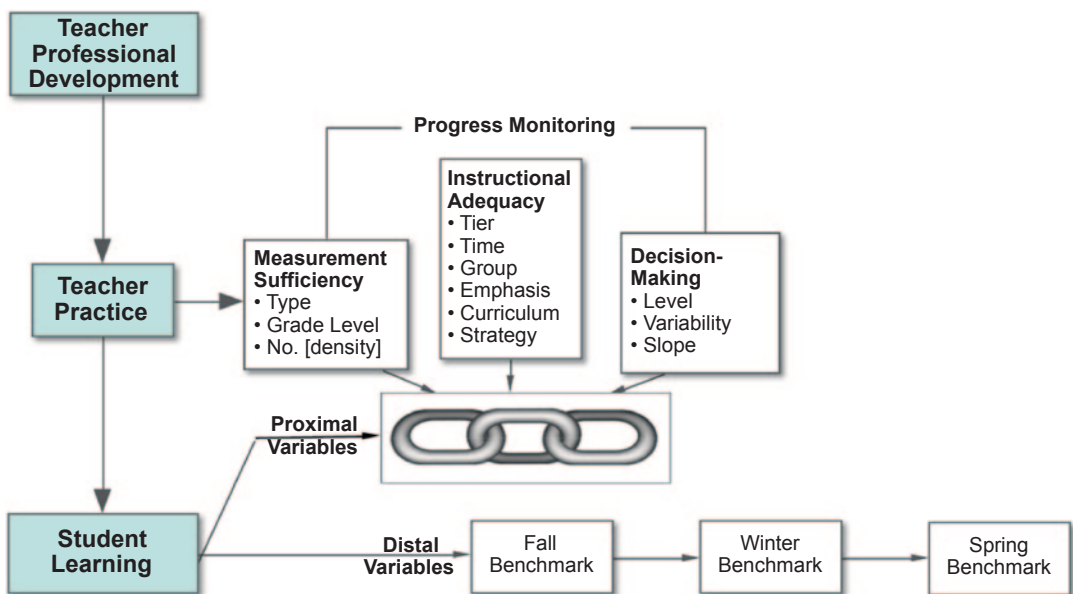


Fig. 9 CBM professional development → teacher change → student learning

Rather, the effect needs to close the achievement gap in which students at risk are catching up to their peers on grade-level performance measures (e.g., benchmarks). Finally, for systemic change, data need to be collected on proximal variables for developing reports on use, allowing for professional development to be tailored and specific.

Professional Development

Few studies have either documented training research on data-based decision-making (Stecker et al. 2005) or investigated effects of decision-making using a single-subject design (Hofstetter 2003). For the most part, training has focused on administration and scoring of measures rather than on how to use the outcomes to define critical components of instruction, integrating diagnosis of instruction based on decision-making that is hypothesis driven.

Stecker et al. (2005) documented investigations ($n=7$) that used “data-based decision rules” (with general outcomes) or used a skill analysis (mastery monitoring). The authors also considered in their summary whether achievement effects were found ($n=4$) or no achievement effects were found ($n=3$). In the end, they concluded training needed to move beyond simple collection of CBM data to affect student achievement. Rather, teachers must be trained to use the CBM measures to evaluate their instructional effectiveness; when student progress is less than expected, teachers then must be trained to make program modifications. They further noted that “raising goals when teachers underestimate student performance also appears to affect student growth. Data-management software also emerged as an aid to teachers in their use of CBM” (p. 803). Finally, when training included skills-analysis information, it “helped teachers examine student performance and highlighted skills for remediation” (p. 807). The authors concluded that not only do teachers need training but also the use of single-subject designs needs to be incorporated into the research on practice.

Hofstetter (2003) used a multiple-baseline (across passages) design with a first-grade student to investigate instruction and the addition of reward conditions to improve a student’s oral reading fluency. Instruction included listening to the passage being read, repeated reading of the passage, and an error correction routine; in addition, a peer-mediated (tutoring) reading session was included. As the author reports, “The remarkable aspect of this finding is that the results were obtained with low-word-overlap passages” (p. 645). What this means is that when teachers are trained in a single-subject design, they can obtain significant results that transfer well beyond the corpus of words being directly taught. Importantly, single-subject design and methodology are the foundation upon which RTI rests, yet the RTI literature is less than rich in empirical examples and often this literature needs to be accessed through more traditional applied behavior analysis publications and search terms (e.g., *Journal of Applied Behavior Analysis* or *Journal of the Experimental Analysis of Behavior* using search phrases such as “brief experimental analysis” or “functional academic assessment”). In summary, an effective technology tool supporting RTI would include training in the systematic use of data, content standards covered, consideration of goals, skills analysis, and data management systems; it also would include use of single-subject designs in practice. However, most RTI training systems have little data on the effectiveness of the training, even though they are premised upon the collection of student performance and progress. Another problem arises when data are collected only at the individual student level and not on teachers. Therefore, professional development and practice cannot be tailored to the areas in which teachers need the most assistance, whether it is about how to effectively progress monitor students, how to develop effective instructional programs, or how to decide on maintaining or changing programs.

The graphic displays on technology-enhanced assessment systems can be designed to structure teachers’ RTI, guiding them to make instructional decisions, while concurrently monitoring student progress. To be most effective for systems-wide

application, such systems should also collect use data, so that policy-makers and administrators can apply the information to fine-tune their practice, improve the RTI system, and develop sensitive policies for practice. Presently, most RTI models operate primarily for individual students and on a *laissez-faire* basis with little feed forward or feedback in which progress is monitored, instruction is developed, or decisions are made. Although some research studies provide evidence of the effectiveness of using student performance data to assist teachers in providing interventions with integrity (see, e.g., VanDerHeyden et al. 2012), such checks and balances are not yet standard in assessment systems, providing rich ground for future development in this area.

The information on instruction through progress monitoring is expected to result in improved outcomes because the entire RTI process can become more systematic with evaluation data for teachers and students. For teachers, the feed forward and feedback features allow them to not only focus on the individual student but also to generalize in a manner that is likely to make their work more efficient. Training can focus on aggregated data within grade levels and measures and then be broken down in various ways within and across the three components of measurement sufficiency, instructional adequacy, and decision-making.

Technology-based assessments can provide teachers and students “next-generation” information through relational databases. For teachers, access can be provided to systematic models of assessment practices with classroom vignettes, exemplary practices, and resources that can be immediately used to develop effective progress monitoring. Student reports on content (standards) coverage as well as other aspects of instruction, combined with progress on both proximal and distal outcomes can be accessed to evaluate not only the effectiveness of interventions but also the systematic use of data-based decision-making. Such information can help teachers focus not only on how to implement best practice but also how to interpret information on student performance and progress. For example, a teacher could determine how many students are

being monitored on specific measures aligned with specific standards, grade levels, and time intervals; how students are being organized into tiers, time, and groups, as well as specific instructional emphases being used (along with curriculum materials and strategies); how well decisions have been made with subsequent changes in level, variability, or slope. But realizing these potential benefits of technology-based assessments requires information and professional development in how to use the information in a systematic manner.

Enhancement to Training Through Effective Reports Future directions for enhancing technology-based assessment training, reports, and support materials (for easyCBM[®] and other learning management systems), should address the following components: measurement sufficiency, instructional adequacy, and decision-making.

- Information on benchmark performance for each of the measures should be analyzed and student performance used to recommend an appropriate measure type and grade level for progress monitoring (measurement sufficiency).
- Information on using the progress-monitoring system for specific students who are at risk and the number of days since the last progress measurement for each type and grade level of measure being used should be computed. This training could be used to select the student and the measure so the teacher can immediately begin measurement with either one-on-one administration (letter names and sounds, phoneme segmentation, and word or passage reading fluency) or group administration (vocabulary, comprehension, mathematics; measurement sufficiency).
- Use of benchmark measurement (and assignment to three levels of risk) can be used to group students by tier and size so teachers can organize groups with known levels of risk or ensure students at risk are distributed appropriately in various groups (instructional adequacy).
- Instruction/professional development should be embedded within the system, covering

such topics as: (a) instructional tier (1–3), (b) allocating time, (c) defining primary instructional emphasis (standards), (d) defining instructional strategy, and (e) providing curriculum and supplemental materials (instructional adequacy).

- Professional development should be included addressing using student progress to evaluate the current instructional program, including how to analyze for each progress measure (a) running average (cumulative level) of performance, (b) average variation (standard deviation), and (c) slope (displaying the average growth per week in the measure; decision-making).
- Annual goals and calculations the discrepancy between current changes and eventual (expected) performance using the three indices above with a “guesstimate” of likelihood (very likely, somewhat likely, and unlikely; decision-making).
- Information that informs decisions to “change or maintain” instruction associated with the number of measures (being used to calculate the progress) and the changeover (decision-making).

Approaching the topic of technology-based assessments by addressing warrants from a validity framework, a markedly different environment from current practice can be provided. First, the information can be integrated, using the power of databases to provide real-time reports drawing from a variety of sources (e.g., student performance data with individual, criterion, and normative referents; teacher instructional practices, including curriculum, strategies, intensity, duration, frequency, etc.; student demographics, enabling analysis of impact by various subgroups). Second, the system can be preventative, in that teachers may go back and forth between reports and actions. Third, such reports can tie together multiple databases for more appropriate reports on systems functioning targeted at professional development. This validity framework approach informs all parts of the discussion related to how the different parts of the technology-based assessment must work in concert to achieve the results sought in adopting an RTI approach: enhanced

teacher understanding of data-based decision-making and improved student learning outcomes.

References

- Alonzo, J., & Tindal, G. (2007). *The development of early literacy measures for use in a progress monitoring assessment system: Letter Names, Letter Sounds and Phoneme Segmenting (Technical Report No. 39)*. Eugene: Behavioral Research and Teaching, University of Oregon.
- Alonzo, J., & Tindal, G. (2009a). *Alternate form and test-retest reliability of easyCBM[®] reading measures (Technical Report No. 0906)*. Eugene: Behavioral Research and Teaching, University of Oregon.
- Alonzo, J., & Tindal, G. (2009c). *The development of K-8 progress monitoring measures in mathematics for use with the 2% and general education populations: Grade 1 (Technical Report No. 0919)*. Eugene: Behavioral Research and Teaching, University of Oregon.
- Alonzo, J., & Tindal, G. (2009b). *The development of K-8 progress monitoring measures in mathematics for use with the 2% and general education populations: Kindergarten (Technical Report No. 0921)*. Eugene: Behavioral Research and Teaching, University of Oregon.
- Alonzo, J., Tindal, G., Ulmer, K., & Glasgow, A. (2006). *easyCBM[®] online progress monitoring assessment system*. Eugene: University of Oregon, Behavioral Research and Teaching. (<http://easyCBM.com>).
- Alonzo, J., Lai, C. F., & Tindal, G. (2009e). *The development of K-8 progress monitoring measures in mathematics for use with the 2% and general education populations: Grade 3 (Technical Report No. 0902)*. Eugene: Behavioral Research and Teaching, University of Oregon.
- Alonzo, J., Lai, C. F., & Tindal, G. (2009d). *The development of K-8 progress monitoring measures in mathematics for use with the 2% and the general education populations: Grade 2 (Technical Report No. 0920)*. Eugene: Behavioral Research and Teaching, University of Oregon.
- Alonzo, J., Lai, C. F., & Tindal, G. (2009f). *Mathematics for Use with the 2% and General Education Populations: Grade 4 (Technical Report No. 0903)*. Eugene: Behavioral Research and Teaching, University of Oregon.
- Alonzo, J., Nese, J. F. T., Lai, C. F., Anderson, D., Park, B. J., & Tindal, G. (2010). *The alignment of easyCBM[®] math measures to curriculum standards (Technical Report No. 1002)*. Eugene: Behavioral Research and Teaching, University of Oregon.
- Alonzo, J., Anderson, D., Park, B. J., & Tindal, G. (2012a). *The development of CBM vocabulary measures: Grade 2 (Technical Report No. 1209)*. Eugene: Behavioral Research and Teaching, University of Oregon.

- Alonzo, J., Anderson, D., Park, B. J., & Tindal, G. (2012b). *The development of CBM vocabulary measures: Grade 3 (Technical Report No. 1210)*. Eugene: Behavioral Research and Teaching, University of Oregon.
- Alonzo, J., Anderson, D., Park, B. J., & Tindal, G. (2012c). *The development of CBM vocabulary measures: Grade 4 (Technical Report No. 1211)*. Eugene: Behavioral Research and Teaching, University of Oregon.
- Alonzo, J., Anderson, D., Park, B. J., & Tindal, G. (2012d). *The development of CBM vocabulary measures: Grade 5 (Technical Report No. 1212)*. Eugene: Behavioral Research and Teaching, University of Oregon.
- Alonzo, J., Anderson, D., Park, B. J., & Tindal, G. (2012e). *The development of CBM vocabulary measures: Grade 6 (Technical Report No. 1213)*. Eugene: Behavioral Research and Teaching, University of Oregon.
- Alonzo, J., Anderson, D., Park, B. J., & Tindal, G. (2012f). *The development of CBM vocabulary measures: Grade 7 (Technical Report No. 1214)*. Eugene: Behavioral Research and Teaching, University of Oregon.
- Alonzo, J., Anderson, D., Park, B. J., & Tindal, G. (2012g). *The development of CBM vocabulary measures: Grade 8 (Technical Report No. 1215)*. Eugene: Behavioral Research and Teaching, University of Oregon.
- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2014). *Standards for educational and psychological testing*. Washington, DC: American Psychological Association.
- Anderson, D., Tindal, G., & Alonzo, J. (2009). *Internal consistency of general outcome measures in grades 1–8 (Technical Report No. 0915)*. Eugene: Behavioral Research and Teaching, University of Oregon.
- Anderson, D., Lai, C. F., Nese, J. F. T., Park, B. J., Sáez, L., Jamgochian, E. M., & Tindal, G. (2010). *Technical adequacy of the easyCBM® primary-level mathematics measures (Grades k-2), 2009–2010 version (Technical Report No. 1006)*. Eugene: Behavioral Research and Teaching, University of Oregon.
- Anderson, D., Alonzo, J., & Tindal, G. (2010a). *easy-CBM® mathematics criterion related validity evidence: Oregon state test (Technical Report No. 1011)*. Eugene: Behavioral Research and Teaching, University of Oregon.
- Anderson, D., Alonzo, J., & Tindal, G. (2010b). *easy-CBM® mathematics criterion related validity evidence: Washington state test (Technical Report No. 1010)*. Eugene: Behavioral Research and Teaching, University of Oregon.
- Anderson, D., Lai, C. F., Alonzo, J., & Tindal, G. (2011). Examining a grade-level math CBM designed for persistently low-performing students. *Educational Assessment, 16*, 15–34. doi:10.1080/10627197.2011.551084.
- Anderson, D., Alonzo, J., & Tindal, G. (2011a). *A cross-validation of easyCBM® mathematics cut scores in Oregon: 2009–2010 (Technical Report No. 1104)*. Eugene: Behavioral Research and Teaching, University of Oregon.
- Anderson, D., Alonzo, J., & Tindal, G. (2011b). *A cross-validation of easyCBM® mathematics cut scores in Washington state: 2009–2010 Test (Technical Report No. 1105)*. Eugene: Behavioral Research and Teaching, University of Oregon.
- Anderson, D., Alonzo, J., & Tindal, G. (2011c). *easy-CBM® reading criterion related to validity evidence: Washington state test 2009–2010 (Technical Report No. 1101)*. Eugene: Behavioral Research and Teaching, University of Oregon.
- Anderson, D., Alonzo, J., & Tindal, G. (2011d). *easy-CBM® reading criterion related validity evidence: Oregon state test 2009–2010 (Technical Report No. 1103)*. Eugene: Behavioral Research and Teaching, University of Oregon.
- Anderson, D., Lai, C., Park, B., Alonzo, J., & Tindal, G. (2012a). *An examination of test-retest, alternate form reliability, and generalizability theory study of the easyCBM® reading assessments: Grade 2 (Technical Report # 1217)*. Eugene: Behavioral Research and Teaching.
- Anderson, D., Park, B. J., Lai, C., Alonzo, J., & Tindal, G. (2012b). *An examination of test-retest, alternate form reliability, and generalizability theory study of the easyCBM® reading assessments: Grade 1 (Technical Report # 1216)*. Eugene: Behavioral Research and Teaching, University of Oregon.
- Embretson, S., & Reise, S. (2000). *Item response theory for psychologists*. Mahwah: Lawrence Erlbaum.
- Hofstetter, C. (2003). Contextual and mathematics accommodation test effects for English-language learners. *Applied Measurement in Education, 16*, 159–188. doi:10.1207/S15324818AME1602_4.
- Irvin, P. S., Park, B. J., Anderson, D., Alonzo, J., & Tindal, G. (2011). *A cross-validation of easyCBM® reading cut scores in Washington: 2009–2010 (Technical Report No. 1109)*. Eugene: Behavioral Research and Teaching, University of Oregon.
- Irvin, P. S., Alonzo, J., Lai, C. F., Park, B. J., & Tindal, G. (2012a). *Analyzing the reliability of the easyCBM® reading comprehension measures: Grade 7 (Technical Report No. 1206)*. Eugene: Behavioral Research and Teaching, University of Oregon.
- Irvin, P. S., Alonzo, J., Park, B. J., Lai, C. F., & Tindal, G. (2012b). *Analyzing the reliability of the easyCBM® reading comprehension measures: Grade 6 (Technical Report No. 1205)*. Eugene: Behavioral Research and Teaching, University of Oregon.
- Lai, C. F., Alonzo, J., & Tindal, G. (2009a). *The development of K-8 progress monitoring measures in mathematics for use with the 2 % and general education populations: Grade 5 (Technical Report No. 0901)*. Eugene: Behavioral Research and Teaching, University of Oregon.

- Lai, C. F., Alonzo, J., & Tindal, G. (2009d). *The development of k-8 progress monitoring measures in mathematics for use with the 2 % and general education populations: Grade 8 (Technical Report No. 0904)*. Eugene: Behavioral Research and Teaching, University of Oregon.
- Lai, C. F., Alonzo, J., & Tindal, G. (2009c). *The development of k-8 progress monitoring measures in mathematics for use with the 2 % and general populations: Grade 7 (Technical Report No. 0908)*. Eugene: Behavioral Research and Teaching, University of Oregon.
- Lai, C. F., Alonzo, J., & Tindal, G. (2009b). *The development of K-8 progress monitoring measures in mathematics for use with the 2 % and the general education populations: Grade 6 (Technical Report No. 0907)*. Eugene: Behavioral Research and Teaching, University of Oregon.
- Lai, C. F., Nese, J. F. T., Jamgochian, E. M., Alonzo, J., & Tindal, G. (2010). *Technical adequacy of the easyCBM[®] primary-level reading measures (grades K-1), 2009–2010 version (Technical Report No. 1003)*. Eugene: Behavioral Research and Teaching, University of Oregon.
- Lai, C. F., Irvin, P. S., Alonzo, J., Park, B. J., & Tindal, G. (2012a). *Analyzing the reliability of the easyCBM[®] reading comprehension measures: Grade 2 (Technical Report No. 1201)*. Eugene: Behavioral Research and Teaching, University of Oregon.
- Lai, C. F., Irvin, P. S., Park, B. J., Alonzo, J., & Tindal, G. (2012b). *Analyzing the reliability of the easyCBM[®] reading comprehension measures: Grade 3 (Technical Report No. 1202)*. Eugene: Behavioral Research and Teaching, University of Oregon.
- Liu, K., Ketterlin-Geller, L. R., Yovanoff, P., & Tindal, G. (2008). *Examining item functioning of math screening measures for grades 1–8 students (Technical Report No. 0804)*. Eugene: Behavioral Research and Teaching, University of Oregon.
- Nese, J. F. T., Lai, C. F., Anderson, D., Jamgochian, E. M., Kamata, A., Saez, L., & Tindal, G. (2010a). *Technical adequacy of the easyCBM[®] mathematics measures: Grades 3–8: 2009–2010 version (Technical Report No. 1007)*. Eugene: Behavioral Research and Teaching, University of Oregon.
- Nese, J. F. T., Lai, C. F., Anderson, D., Park, B. J., Tindal, G., & Alonzo, J. (2010b). *The alignment of easyCBM[®] math measures to curriculum standards (Technical Report No. 1002)*. Eugene: Behavioral Research and Teaching, University of Oregon.
- Park, B. J., Irvin, P. S., Lai, C. F., Alonzo, J., & Tindal, G. (2012a). *Analyzing the reliability of the easyCBM[®] reading comprehension measures: Grade 5 (Technical Report No. 1204)*. Eugene: Behavioral Research and Teaching, University of Oregon.
- Park, B. J., Irvin, P. S., Alonzo, J., Lai, C. F., & Tindal, G. (2012b). *Analyzing the reliability of the easyCBM[®] reading comprehension measures: Grade 4 (Technical Report No. 1203)*. Eugene: Behavioral Research and Teaching, University of Oregon.
- Stecker, P., Fuchs, L., & Fuchs, D. (2005). Using curriculum-based measurement to improve student achievement: Review of research. *Psychology in the Schools*, 42, 795–819. doi:10.1002/pits.20113.
- Thompson, S. J., Johnstone, C. J., Anderson, M. E., & Miller, N. A. (2005). *Considerations for the development and review of universally designed assessments (Technical Report 42)*. Minneapolis: University of Minnesota, National Center on Educational Outcomes. <http://education.umn.edu/NCEO/OnlinePubs/Technical42.htm>.
- Tindal, G. (2013). Curriculum-based measurement: A brief history of nearly everything from the 1970s to the present. *ISRN Education*, 2013, 1–29. doi:10.1155/2013/958530.
- Tindal, G., Alonzo, J., & Anderson, D. (2009a). *Local normative data on easyCBM[®] reading and mathematics: Fall 2009 (Technical Report No. 0918)*. Eugene: Behavioral Research and Teaching, University of Oregon.
- VanDerHeyden, A. M., McLaughlin, T., Algina, J., & Snyder, P. (2012). Randomized evaluation of a supplemental grade-wide mathematics intervention. *American Education Research Journal*, 49, 1251–1284.