

Best Practices in Using Curriculum-Based Measurement in a Problem-Solving Model

Mark R. Shinn

National Louis University

OVERVIEW

Curriculum-based measurement (CBM) is a set of standardized and short duration tests (i.e., 1–5 minutes) used by educators to evaluate the effects of their instructional interventions in the basic skills of reading, mathematics, spelling, and written expression (Shinn, 1989, 1998). For example, important decisions about general reading skill are made by having a student read a passage aloud for 1 minute, counting the number of words read correctly (WRC).

CBM was developed by Stanley Deno and others at the University of Minnesota Institute for Research on Learning Disabilities more than 25 years ago to give teachers simple tools to write Individualized Educational Program (IEP) goals and monitor progress (Deno, 1985, 1986, 1995, 2002, 2003; Deno & Mirkin, 1980). However, its use beyond special education quickly expanded to give *all* educators simple, accurate tools to scale achievement for both universal screening and progress monitoring within a problem-solving model (Deno, 1995, 2002, 2003; Germann & Tindal, 1985; Marston & Magnusson, 1985; Tindal, Wesson, Deno, Germann, & Mirkin, 1985). More than a quarter century of scientific evidence has accumulated demonstrating that CBM provides reliable and valid measures of *general achievement* (e.g., general reading achievement) that is sensitive to student improvement and that improves student achievement when used to monitor progress (Deno, Fuchs, Marston, & Shin, 2001; Fuchs & Fuchs, 1986, 1988, 1999, 2004).

Historically, many school psychologists have not been interested in, nor collected, CBM progress-monitoring data. The perception was that progress monitoring is a

teacher's responsibility. However, in earlier editions of *Best Practices*, I suggested that although school psychologists typically do not *collect* CBM progress monitoring data, school psychologists should still be interested in and *use* CBM in decision making (see Shinn, 1995, 2002, for more detail). I offered as a rationale that school psychologists should be concerned about the effectiveness of special education programs for students with disabilities and the educational and psychological needs of *all* children. School psychologists have had a long-standing concern about the effectiveness of special education, and these opinions are echoed by every meta-analysis of special education effectiveness since 1980 (Carlberg & Kavale, 1980; Kavale & Forness, 1985, 1999). Additionally, many school psychologists are interested in providing instructional and behavioral consultation and support to general education teachers and parents to promote healthy development (Ikeda et al., 2002; Ikeda, Tilly, Stumme, Volmer, & Allison, 1996).

These concerns have not diminished, but increased, in visibility in (a) educational science and practice, and (b) educational policy and law. In particular, No Child Left Behind (NCLB) and the 2004 reauthorization of the Individuals with Disabilities Education Act (IDEA) make clear the importance of CBM as an integral tool in evidence-based practices. The importance of tools like CBM are now expressed explicitly in *School Psychology: A Blueprint for Training and Practice III* (Ysseldyke et al., 2006) competencies for (a) data-based decision making and accountability, (b) technological applications, and (c) application of science and the scientific method (see Ysseldyke et al., chapter 3, vol. 1, for more detail).

This chapter briefly reviews developments in educational science and practice and educational policy and law that make school psychologists' use of CBM in a problem-solving model vital (Deno, 1989, 1995, 2002). The chapter then describes the key features of CBM and provides illustrations of how CBM is used in a three-tiered problem-solving model, including response to intervention (RTI).

BASIC CONSIDERATIONS

In its early form (i.e., 1979–1995), CBM testing materials typically were derived directly from students' general education curriculum. If the general education reading curriculum was Reading Mastery, then the reading probes were constructed from Reading Mastery passages. When testing materials were drawn directly from general curriculum (i.e., curriculum-specific measurement), the assessment outcomes had high instructional validity (Fuchs & Deno, 1992, 1994). Educators knew which words a student read correctly. However, curriculum-specific testing materials had a number of technical and logistic liabilities. With respect to technical liabilities, because the general curriculum source material often varied in difficulty within any given level, passage difficulty became a source of error variability and resulted in a loss of accuracy in depicting student progress (Fuchs & Deno, 1994; Hintze, Shapiro, & Daly, 1998; Hintze, Shapiro, & Lutz, 1994). From a logistics perspective, teachers spent time developing test materials, some schools did not use general curriculum, general curriculum varied school to school or class to class, or general curriculum changed with sufficient frequency that the test development process started over. As a result, the development of a database that was consistent across settings and time was hampered.

Educational Science and Practice

Advances in science demonstrated that CBM's most important feature was *not* that the test materials come from a specific curriculum, but that the testing process is standardized and test materials are of equal difficulty and represent the general curriculum (Fuchs & Deno, 1994; Hintze et al., 1998; Hintze et al., 1994). The first example of standardized, validated general outcome CBM reading probes was the Test of Oral Reading Fluency (Deno, Deno, Marston, & Marston, 1987). Once the test construction process became focused on the development of general curriculum test materials of equal difficulty, it was possible to evaluate and report

alternate form reliability and passage difficulty and produce technical manuals (e.g., Howe & Shinn, 2002). Publishers of CBM assessment materials have proliferated in recent years (e.g., AIMSweb, Dynamic Indicators of Basic Early Literacy Skills [DIBELS], Edcheckup). Importantly, the use of standard general curriculum probes eliminated the logistics issues identified earlier and allowed the expansion of CBM to other important decisions about students, notably universal screening (Shinn, Shinn, Hamilton, & Clarke, 2002) and allowing evaluation of student progress toward passing high-stakes accountability tests (Hintze & Silbergliitt, 2005; Silbergliitt & Hintze, 2005).

Educational Policies and Law

As noted by Tilly (chapter 2, vol. 1), “NCLB has created a new context. Since 2002, schools have become avid collectors and users of data.” The themes reflected in NCLB, evidence-based interventions, early intervention, universal screening, and data-based decision making and progress monitoring are also echoed in IDEA 2004. IDEA, in brief, specifies that local education agencies can use a process of RTI to determine students' eligibility for special education under the category of learning disabilities and not use the common method of computing the ability–achievement discrepancy.

At the heart of the RTI process is the dual discrepancy (Batsche et al., 2005; Fuchs, Fuchs, & Speece, 2002; Pericola Case, Speece, & Eddy Molloy, 2003), where students must (a) be significantly different from their peers in educational achievement *and* (b) not be improving at an adequate rate when given high quality, scientifically based instruction. IDEA 2004 provides school psychologists an opportunity to no longer serve as the special education gatekeepers through use of the ability–achievement discrepancy and allows them to contribute to scientifically based instruction and interventions. Of course, because CBM has been used to answer these questions for more than 25 years, knowledge of this assessment technology is now critical to school psychologists.

CBM is also valuable to remedy other legal concerns. RTI has dramatically increased interest in CBM as a progress-monitoring tool for those students considered for special education eligibility because it provides a scientifically based method of determining adequate progress. Ironically, an objective method with scientifically based goals and progress-monitoring methods has remained elusive for students *who receive special education*.

In the early 1990s, concern was expressed around the quality of IEP goals (Smith, 1990). After almost another 10 years, concern about IEP goal quality was expressed again (Bateman & Linden, 1998).

As discussed in more detail in Shinn and Shinn (2000), the U.S. Congress set forth to remedy this issue in the 1997 revision of IDEA. Explicitly stated within this revision was the need to (a) assess educational need; (b) write measurable annual IEP goals, monitor progress, and report progress to parents at least as often as progress is reported about nondisabled students; and (c) make revisions in the IEP to address any unexpected lack of progress. The critical component in each of the explicit statements is the need to make these decisions and actions in the context of participation and progress in *general curriculum*. The ability of CBM to improve the IEP annual goals and progress monitoring intents of recent federal laws for those students who receive special education is another reason that it should be of interest to school psychologists and educators alike.

The national needs for scientifically based progress monitoring using CBM has resulted in a major policy agenda to support implementation in schools. In 2004, the U.S. Department of Education, Office of Special Education Programs (OSEP), funded the National Student Progress Monitoring Center for 5 years (www.studentprogress.org). The center's goals include the following:

- Raise knowledge and awareness by forming partnerships and communicating with states, districts, associations, technical assistance providers, institutions of higher education, and other interested groups
- Provide implementation support for using and sustaining proven progress monitoring practices to states and districts
- Provide for national dissemination by developing resources and supporting ongoing information sharing through advanced web services, regional meetings, and at national conferences

Among the center's most important tasks is the identification of tools that meet the standards for scientifically based progress monitoring. A review of the center's analysis shows that most of the tools that meet standards are based on CBM. The center's dissemination efforts are also augmented by OSEP's establishment of three CBM national demonstration and dissemination centers at the University of Minnesota, Lehigh University, and the University of Oregon.

What Defines CBM

Since 2000, practitioners have encountered a virtual explosion of information in the professional literature on a wide range of testing strategies that they can use for scientifically based progress monitoring and universal screening in a problem-solving model:

- Is curriculum-based assessment (CBA; Shapiro, 2004) the same thing as CBM?
- Is DIBELS (Kaminski & Good, 1996) the same or different from CBM?

Compounding the confusion are other techniques that use short duration tests and/or oral reading (e.g., informal reading inventories) or measure early literacy, and its authors claim they can be used as progress-monitoring measures.

In brief, all of these measures, including CBM, fit under the general label of CBA. This term represents *any* testing strategy that uses a student's curriculum, general or specific, as the basis for decision making. Unfortunately, CBA has become so broad that the descriptor is more of a liability for understanding than an asset. (For more information on distinctions among types of CBM, see Shinn, Rosenfield, & Knutson, 1989; Shinn & Bamonto, 1998; Tindal, 1993.)

CBM is a particular *type* of CBA. Testing is accomplished by using a limited number of measures, albeit standardized and validated, of student performance in the basic skill areas of reading, spelling, mathematics computation and application, and written expression, early literacy and numeracy. Members of the CBM "family" are constructed and validated using criteria specified originally by Jenkins, Deno, and Mirkin (1979) and later refined by Fuchs and Fuchs (1999). Some of these criteria include standardized testing, reliability, validity, sensitivity to improvement, and high efficiency, including ease of training and administration and scoring. Thus, DIBELS is a *type* of CBM and, in fact, was designed to be a downward extension of CBM reading, allowing the early identification and progress monitoring of Kindergarten and Grade 1 at risk student (Kaminski & Good, 1998).

Core CBM Tools

CBM consists of the following foundational tests:

- *Reading CBM (R-CBM)*: Students read aloud from text for 1 minute. The number of words read correctly constitutes the basic decision-making metric (Deno,

- Mirkin, & Chiang, 1982; Fuchs, Fuchs, Hosp, & Jenkins, 2001; Miura Wayman, Wallace, Ives Wiley, Ticha, & Espin, 2007; Shinn & Shinn, 2003a).
- *Maze-CBM*: Maze is a multiple choice cloze reading technique. Students read silently and select the correct word that preserves meaning from three choices when every seventh word is deleted (Fuchs & Fuchs, 1992; Shinn & Shinn, 2003b). The number of correct word choices per 3 minutes is the primary metric.
 - *Spelling CBM*: Students write randomly selected words from a pool of grade-level words (e.g., Shinn & Shinn, 2003c) that are dictated orally at specified intervals (either 5, 7, or 10 seconds) for 2 minutes. The number of correct letter sequences and words spelled correctly are counted (Fuchs, Fuchs, Hamlett, & Allinder, 1991).
 - *Written expression CBM*: Students write a story for 3 minutes after being given a story starter from a pool of age-appropriate story starters (Powell-Smith & Shinn, 2004). The number of words written, spelled correctly, and/or correct word sequences is counted (Deno, Marston, & Mirkin, 1982; Marston, 1989; McMaster & Espin, 2007).
 - *Mathematics computation CBM*: Students write answers to grade-level computational problems (Shinn & Shinn, 2004) via 2- to 4-minute probes. The number of correctly written digits is counted (Foegen, Jiban, & Deno, 2007).
 - *Math applications CBM*: Students write answers to grade-level mathematics application problems on 4-minute probes (Fuchs, Fuchs, & Hamlett, 1995). The number of correct problems is counted (Foegen et al., 2007; Fuchs et al. 1994)

Dynamic Indicators of Basic Skills as a Foundational CBM Concept

Where do these simple and short tests fit in an era where test construction has become highly sophisticated and students take a variety of significantly longer and varied tests each day? In the mid-1980s, CBM researchers and trainers gathered to find a simple way to communicate the intent of CBM because it was so different from standard educational practice. The outcome of the discussion was the big idea that CBM can be represented by the mnemonic dynamic indicators of basic skills (DIBS). The *B* and the *S* correspond to the broad academic domains vital to school success, basic skills. In fact, when DIBELS was developed years later, it expanded on the DIBS pneumonic and replaced basic skills with early literacy skills.

CBM is *dynamic* in that the measures are designed to be sensitive to the short-term effects (i.e., 4–6 weeks) of instructional interventions. They are designed to assess *change*. Because they are sensitive to improvement, they make excellent tools for progress monitoring, whether it is required by an IEP annual goal, a short-term goal as part of response to intervention (RTI), or, as will be shown, as part of progress monitoring of *all* general education students. Because the tests are short (i.e., 1–4 minutes) in addition to being sensitive to improvement, they can be administered frequently, even one to two times per week, to allow learning to be assessed on a routine basis without a significant loss of instructional time or personnel resources.

The *I* in DIBS is the most controversial. This part of the mnemonic represents the big idea that the CBM measures are designed to serve as *indicators* or constructs of an achievement domain. Some professionals judge a test's validity based on opinion as to what a test measures. For some practitioners, R-CBM, where students read aloud for 1 minute, the construct measured is decoding (Hamilton & Shinn, 2003). However, this judgment is invalid when placed in the context of accumulated construct validity evidence. The goal is to obtain evidence that the behavior sampled by the test (e.g., oral reading, writing spelling words, writing answers to mathematics computation problems) validly represents broader achievement domains (e.g., general reading, spelling, written expression, respectively).

As indicators, CBM measures are *not* designed to represent *all* behaviors that may be included in an academic domain such as reading. For example, R-CBM will not provide direct information about whether a student can separate fact from fiction, or identify how compelling is an author's argument. Furthermore, CBM also does not mean that these tests are the *only* tests used to assess students' educational needs. For example, should a teacher be interested in a student's fact versus fiction skills, curriculum-based evaluation (Howell, Kurns, & Antil, 2002; Howell & Nolet, 1999) would be a suitable assessment process to address this issue.

The selection of the specific CBM test behaviors, scoring methods, and test durations is the result of an ongoing program of research that demonstrates its usefulness in making problem-solving decisions. Beginning in 1978, an extensive program of federally funded research was undertaken to identify key behaviors that would meet the technical requirements to serve as simple, but general, measures of academic performance.

Although a considerable body of information has been published on the technical adequacy of CBM across a number of academic areas, most research has been conducted on R-CBM. (See Marston, 1989, Good & Jefferson, 1998, and miura Wayman et al., 2007; for extensive summaries.) Results suggest that for most elementary-age students, counting the number of words a student reads aloud correctly from connected text in 1 minute, works extremely well as an indicator of general reading proficiency. Evidence has been garnered with respect to construct validity (Fuchs, Fuchs, Hosp, & Jenkins, 2001; Hosp, 2005; Shinn, Good, Knutson, Tilly, & Collins, 1992) as well as traditional conceptions of criterion-related validity, including correlations with other accepted published measures (e.g., Deno, Mirkin, et al., 1982; Fuchs, Fuchs, & Maxwell, 1988) and teacher ratings (Fuchs, Fuchs, & Deno, 1982).

Arguably, the best metaphor to help understand the use of CBM is to think of the CBM measures as educational thermometers. When CBM is used in a problem-solving model, its use as an indicator allows for decisions like those made with a thermometer. Of note also is that like medical thermometers, CBM was not designed to be diagnostic or prescriptive. Like a thermometer that does not inform a physician what caused the high fever or how to treat it, CBM also has these limitations. CBM typically does not tell why a person has an academic performance discrepancy or how to intervene to reduce it. However, this lack of diagnostic and prescriptive efficacy does not diminish CBM's capacity to make other key decisions in an accurate and efficient way.

In sum, CBM's short but repeatable measures from general curriculum were designed to fill a void in educational measurement, especially for students with severe educational needs. They were designed as DIBS to give educators a simple, straightforward way of determining the status of students' educational health in critical academic areas and to provide a mechanism for evaluating the effects of efforts to improve any given student's academic health.

BEST PRACTICES

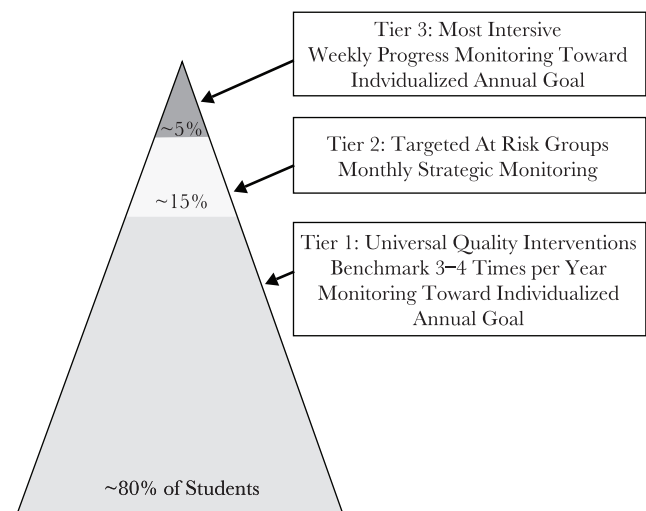
Historically, CBM has been used in *individual* student problem solving at the point of referral to make decisions about potential educational need (problem identification); special eligibility (problem certification); intervention planning and IEP goal setting (exploring solutions); progress monitoring (evaluating solutions);

and periodic, annual, and special education reevaluations (problem solution). However, in the intervening years, the use of CBM as part of proactive problem solving in a three-tier model has become best practice.

Different educational programs historically have had different accountability systems for progress monitoring. For example, general education teachers may use curriculum-embedded tests (e.g., end-of-unit tests) to evaluate reading progress, remedial programs such as Title I may use informal reading inventories, and special education programs may use individually administered, broad-band achievement tests like the Woodcock-Johnson. These different systems have led to a haphazard, inconsistent, and inefficient system of progress monitoring (Shinn et al., 2002). Most importantly, the types of progress-monitoring tools used across different types of programs do not meet standards of scientifically based progress monitoring. Therefore, pictures of a student's progress are difficult to obtain in a system whose primary mission is to produce student learning. The problem of poor progress-monitoring practices is compounded by a lack of regular, routine, and efficient universal screening practices that promote early identification.

CBM, when implemented in a three-tier problem-solving model, provides a uniform, efficient, accurate, and scientifically based tool to conduct both universal screening and progress monitoring across types of instructional programs. The types of CBM progress-monitoring approaches employed in a three-tiered model are shown in Figure 1.

Figure 1. Scientifically based progress monitoring using CBM in a three-tier problem-solving model.



Benchmark Assessment

In Tier 1, general education engages in a process called benchmark assessment where *all* students are tested typically three times per year. In reading, using R-CBM, students are tested individually by reading three grade-level passages from a pool of standardized, field-tested probes (Shinn et al., 2002; Shinn 2003b). This process takes approximately 5 minutes per student per benchmark, or a total of 15 minutes per year, and can be used for both universal screening and progress monitoring. An example of an R-CBM benchmark is shown in Figure 2.

This box-and-whisker chart shows how Arianna, a second grader, performed on grade 2 R-CBM probes at the fall, winter, and spring benchmarks compared to other second graders. R-CBM scores (i.e., number of words read correctly) in the box correspond to the average range (25th–75th percentile). Scores in the upper whisker are those of above average readers (75th–90th percentile). Scores in the bottom whisker (10th–25th percentile) are scores of below average readers and suggest an at-risk status. Scores above and below the whisker are scores above the 90th and below the 10th percentiles, respectively. Arianna’s fall benchmark score of 24 WRC, at the 20th percentile, suggested that she was at risk. Using these data, her classroom teacher ensured she was placed appropriately in a quality core

reading program with flexible skills groupings. Based on her subsequent benchmark scores, Arianna appeared to benefit from her grade 2 reading program as she improved at a faster rate than her peers and reduced the gap, performing in the average range by the end of the year.

CBM is now used proactively in benchmark for two fundamental purposes: (a) universal screening, to identify students at risk for academic failure or with potentially severe educational need, and (b) progress monitoring, to judge educational benefit. Once general education acquires and uses these foundational concepts in everyday decision making, interventions can be made with any student who has an educational need and is not benefiting from intervention, including those students who may need special education. These two concepts form the basis for decision making in an RTI model (Batsche et al., 2005; Fuchs et al., 2002; Pericola Case et al., 2003; Shinn, 2005b).

In a problem-solving model, severe educational need is measured by the performance discrepancy, the difference between how an individual student performs compared with a comparison group (e.g., other students in the school or community) or an empirically defined standard (e.g., low likelihood of passing future high-stakes tests; for more detail, see Shinn, 2005b). R-CBM benchmark results for three students are shown in Figure 3.

Figure 2. Results of grade 2 fall, winter, and spring benchmark assessment for Arianna showing important student progress.

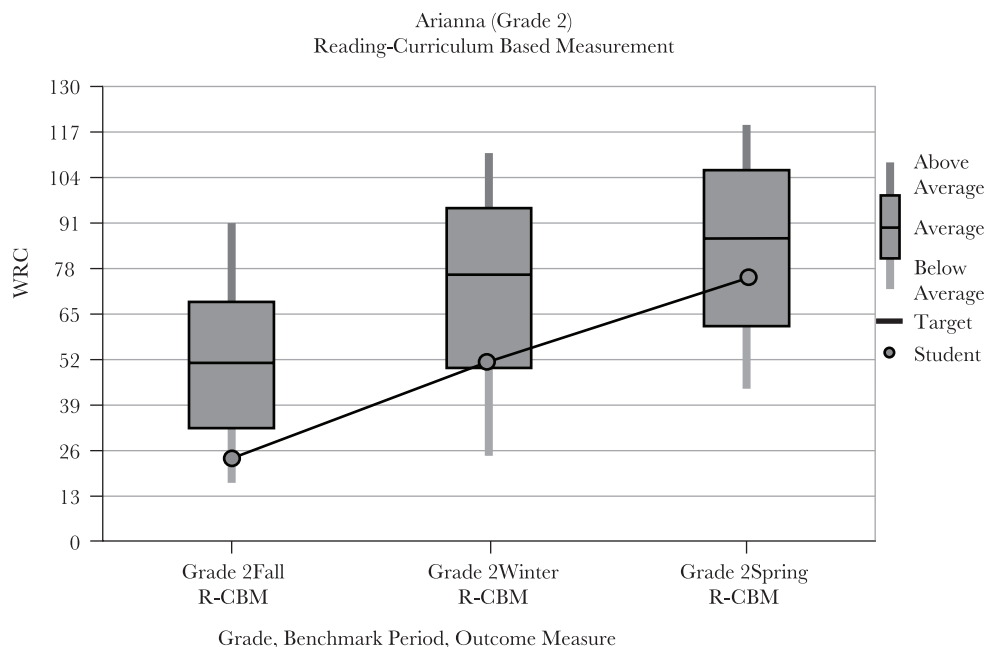
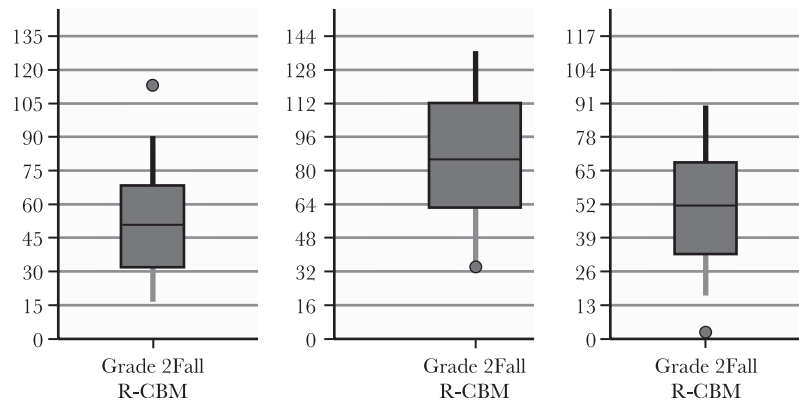


Figure 3. Results of benchmark universal screening showing students with no severe educational need, at risk, and potentially severe educational need.



The first student performs above the 90th percentile and shows no evidence of severe educational need. The second student scores between the 10th and 25th percentile and may be at risk for reading failure. This student would be a likely candidate for Tier 2 intervention and program modification. The third student reads well below the 10th percentile, and if there is no evidence that the student is benefiting from the interventions they are receiving (e.g., Tier 2), then that student may be a good candidate for Tier 3 intervention and individual problem solving.

In a problem-solving model, educational benefit is measured by rate of improvement, compared with the rates of improvement of a comparison group (e.g., other students in the school or community) or changes in likelihood of attaining an empirically defined standard (e.g., low likelihood of passing future high-stakes tests; for more detail, see Shinn, 2005b). A comparison of two students' educational benefit is shown in Figure 4.

The first student's rate of progress was negligible or slightly negative from the fall to winter benchmark. At the fall benchmark, the student read within the range of average readers. By the winter benchmark, the student is now scoring in the range of students considered to be at risk. For this student, the data indicate the need to change the general education reading intervention. The second student's rate of improvement suggests benefit from the reading intervention. The rate of improvement is such that although the student is still well below the level of typically achieving students, the student is reducing the gap from peers. This intervention would likely be continued.

Strategic Monitoring

Benchmark assessment provides an opportunity for teachers to make program or intervention changes at the beginning of the year and at mid-year. For students at risk for failure, educators need more frequent opportunities to make changes if student growth is not what is expected. At Tier 2, this typically takes the form

Figure 4. Results of benchmark testing for two students, only one of whom is benefiting.

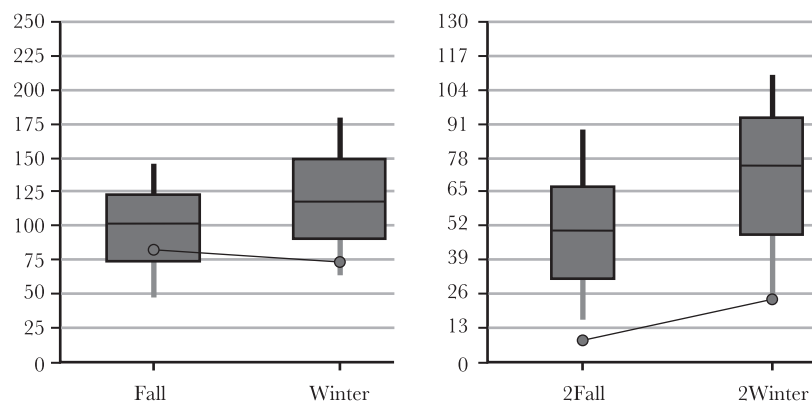
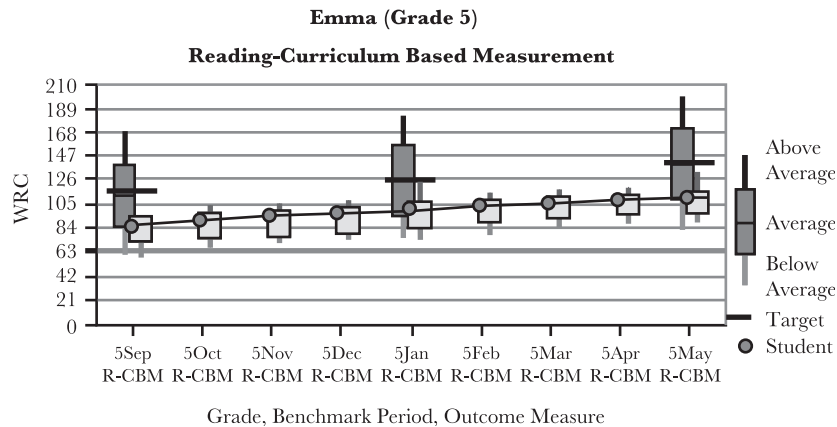


Figure 5. Results of grade 5 strategic monitoring for Emma showing student progress relative to other at-risk students and all grade 5 students in her community.



of repeating the benchmark assessments monthly. An example of strategic monitoring is shown in Figure 5.

Emma, a fifth grader, performed in the range of at-risk students at the fall benchmark and entered the school's Title I program, a Tier 2 intervention. Emma's scores in words read correctly are represented by the individual points on the graph. The smaller box-and-whisker chart is the performance of all at risk students disaggregated from the whole school's benchmark testing fall, winter, and spring or based on additional monthly testing. In the months between benchmarks, students who receive Tier 2 programs are tested using benchmark procedures where they read three grade-level passages. This process takes approximately 5 minutes per student per month and adds approximately six more assessments to the progress-monitoring process than benchmark assessment. Emma's rate of progress was equal to other students who received the Title I intervention. However, if the Title I program was intended to reduce the gap for at-risk students, it is not doing so. Results suggest that the Title I program's interventions should be strengthened.

Frequent Progress Monitoring

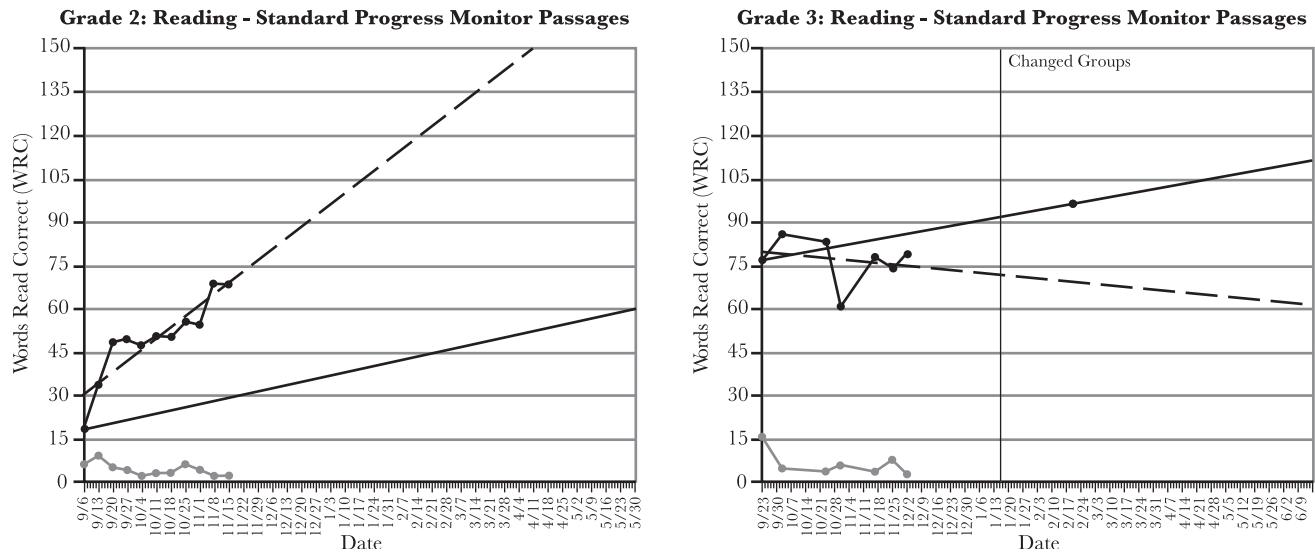
In Tier 3, students either have well-established educational needs (i.e., significant performance discrepancies) or they are at-risk students for whom Tier 2 interventions have not been effective. These students require the most intensive interventions and the most frequent progress monitoring so that effective interventions can be validated and ineffective interventions modified as soon as possible. Student progress monitoring at Tier 3 is characterized by (a) individualized goals and (b) progress monitoring that occurs one to two times

per week. Examples of two students' frequent progress monitoring at Tier 3 using R-CBM are shown in Figure 6.

When a student needs intensive intervention, an individualized goal is written that is to be attained within a specified time frame (e.g., 1 year, 6 weeks). The goal prescribes an expected rate of progress and is shown in both graphs by a solid line. With R-CBM, a student is tested one to two times per week with a single, but different, passage each time. A trend line, or actual rate of improvement, computed using ordinary least squares (Good & Shinn, 1990; Shinn, Good, & Stein, 1989) is shown by the dashed line. For the first student, the actual rate of improvement exceeds the expected rate of improvement. This intervention is effective and, in practice, would trigger decisions about raising the goal or, if the problem is no longer severe, discontinuing the intensive intervention. For the second student, the actual rate of improvement is below the expected rate of improvement. This intervention is *not* effective and, in practice, would trigger decisions about changing the intervention. More detail on frequent progress monitoring will be presented under individual problem-solving activities later in this chapter.

Most students who require this progress-monitoring intensity are students who receive special education or are engaging in an RTI process to determine need for special education. CBM has been validated as an effective tool for goal setting and progress monitoring and for improving achievement for students who receive special education. Numerous studies by Lynn Fuchs and Doug Fuchs (e.g., Fuchs, 1998; Fuchs, Deno, & Mirkin, 1984; Fuchs, Fuchs, & Hamlett, 1989; Fuchs, Fuchs, Hamlett, & Allinder, 1991; Fuchs, Fuchs, Hamlett, & Ferguson, 1992; Fuchs, Fuchs, Hamlett, & Stecker,

Figure 6. Frequent progress monitoring of two students at Tier 3 toward annual IEP goals.



1991) have shown that when teachers use CBM to write data-based goals, monitor the effects of their instructional programs, and adjust their interventions when the data show few effects, student achievement improves. Typical effect sizes—the amount of standard deviation gains in achievement attained by members of the treatment group versus controls—exceed .70 (Fuchs, 1986; Fuchs & Fuchs, 2004).

Individual Problem Solving Using CBM

By using CBM in a three-tier problem-solving model, all educators have access to continuous information regarding students' *educational need* and their *educational benefit* from a range of interventions. By using CBM for universal screening, at-risk students can be identified and placed in Tier 2 and Tier 3 interventions based on benchmark testing rather than waiting for individual students to be referred. Using progress monitoring, an educator can identify which interventions are working or which need improvement.

Many schools, however, have yet to undertake the process of building a data-driven, three-tier problem-solving service delivery system. For these schools, problem solving takes place on a student-by-student basis driven by referral. Additionally, when schools engage in an RTI process, decisions also are made on a student-by-student basis. The remaining portion of this chapter details how CBM data drives problem identification, problem certification, and exploring and evaluating solutions decisions with these students, including RTI.

Problem Identification Decisions

The first step of the problem-solving model is problem identification. The goal of problem identification is to determine in a systematic way if an academic problem exists that is important enough to warrant further assessment (Deno, 1989, 2003, 2005). The core concept in this decision is educational need as measured by the performance discrepancy. If a student has a significant performance discrepancy (i.e., is significantly different in achievement from peers), then additional problem solving is appropriate.

If a school engages in benchmark testing, all general education teachers have access to each student's performance discrepancy datum and can trigger a problem-solving referral for students for whom they have concerns. In an RTI process, a severe performance discrepancy is necessary, but not sufficient for the special education entitlement decision.

What constitutes a potentially severe educational need? In a problem-solving model, problems are defined *situationally* as the discrepancy between what is expected and what occurs within a specific environment or situation. It is important to recognize that the magnitude of the discrepancy remains a value judgment (for more detail, see Shinn, Good, & Parker, 1999). For students identified as learning disabled (LD), converging evidence demonstrates that the defining feature is severe *low* achievement (Fuchs, Fuchs, Mathes, Lipsey, & Eaton, 2000; Fuchs, Fuchs, Mathes, Lipsey, & Roberts, 2001; Gottlieb, Alter, Gottlieb, & Wishner, 1994; Gresham, MacMillan, & Bocian, 1996; Peterson & Shinn, 2002; Reynolds & Heistad, 1997; Shinn,

Ysseldyke, Deno, & Tindal, 1986) compared with typically achieving students and low achievers. For example, a meta-analysis of 79 studies comparing LD and low-achieving students (Fuchs et al., 2000; Fuchs et al., 2001) resulted in an effect size of 0.61. Students placed in special education programs for learning disabilities performed more than one-half standard deviation *lower* in achievement than their low-achieving counterparts.

Since the 1980s, CBM has been used to understand the academic performance characteristics of those students that schools label as LD. Shinn, Tindal, Spira, and Marston (1987) observed that grades 1–6 LD students' R-CBM scores were at the 3rd percentile compared to their general education peers and 75% performed below the 5th percentile. More recently, Peterson and Shinn (2002) found that school-identified LD students from high- and low-achieving communities were about 2 standard deviations below their general education counterparts. Therefore, it is both empirical and logical that a defensible criterion for potential severe educational need is a score below the 10th percentile. Additionally, a more restrictive criterion such as the 5th–7th percentile is defensible based on the published research.

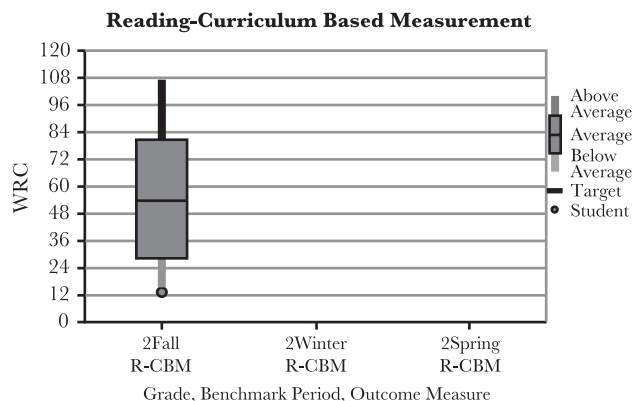
If schools are not using CBM in general education promotion and prevention activities, then local norms must be developed for problem solving. Depending on the degree of school system commitment to a problem-solving model, local norms can be derived from students in the same classroom, building, or district.

For illustration of the problem identification procedures, consider the case of Georges, a second-grade student who was referred to the building problem-solving team (PST) by his general education teacher during the fall of second grade because of concerns about his reading. His scores on grade 2 R-CBM are shown in Figure 7.

If Georges' school engaged in benchmark testing, then a box-and-whisker chart such as the one used in Figure 7 would be used as part of the referral process. If the school did not do benchmarking, Georges would be tested by a member of the PST. Georges' median score on three grade 2 passages was 13 WRC. If his scores were consistent, additional testing for problem identification may not be necessary. If there were high variability (e.g., differences of 30–40 WRC between passages), then Georges would be tested again on another day with three more grade 2 passages.

If there were local norms, Georges' score of 13 WRC would be compared with the range of scores of local

Figure 7. Results of Georges' grade 2 benchmark testing showing potential severe educational need.



peers. If there were no local norms, extensive normative data bases are now available at the state level (e.g., AIMSweb) or from students across the United States (e.g., AIMSweb, DIBELS). In this instance, Georges' school did not have local norms, so his scores were compared with other students in his state. His normative performance was at the 8th percentile, below his school's problem identification standard of the 10th percentile. At this stage of the problem-solving process, the PST ruled out obvious reasons for the teacher's concerns (e.g., poor school attendance, vision or hearing difficulties) and began a process of evaluating his RTI.

Problem Certification

After problem identification (i.e., a severe educational need or a performance discrepancy), a problem certification or special education entitlement decision usually follows. The major activity is to determine the magnitude or *severity* of the problem and whether the student responds to high quality interventions. If a student has a severe educational need and *benefits* from a high-quality intervention in general education, there is no need for a special education entitlement decision. If a student has a severe educational need and *does not benefit* from a high-quality intervention in general education, then more intensive resources such as special education may be required to enable the student to benefit. Only by monitoring progress can the decision of benefit be made.

Determining Severity of Educational Need

The problem certification process begins by conducting a survey-level assessment (SLA) using CBM for each

academic area in which a problem has been identified. In SLA, the student is tested in successively lower levels of the general curriculum. For example, in Georges' case, he was tested in grade 1 and 2 reading using R-CBM. Three passages are administered for each grade, and the median score is calculated. The results of Georges' SLA are presented in Figure 8 with a star used to show his median score. The solid black line through the star allows an interpretation of the R-CBM score relative to normative performance across three benchmark periods per grade. Students are tested in successive levels of R-CBM probes until their performance falls within the average range of peers. This across-grade performance discrepancy facilitates understanding the severity of the problem.

Although Georges read below the 10th percentile compared with grade 2 students, his grade 1 R-CBM scores placed him in the average range for winter grade 1 students. These data show a performance discrepancy compared with same-grade peers but one that may be remediable with high-quality general education instruction. In this instance, the PST decided to determine Georges' response to a more intensive general education intervention. If his R-CBM scores failed to equal the reading skills of grade 1 at any benchmark period, the problem would be considered more severe and his SLA may include performance on other CBM early literacy measures such as Nonsense Word Fluency or Highly Decodable Text.

Some students' needs are so severe that the question of whether a student responds to high-quality interventions in general education can be answered with data-based professional judgment without testing the effects of the intervention. See, for example, the results obtained from an SLA from Ginny, a sixth-grade

student (Figure 9). She read three passages beginning at grade 6, and only when she read grade 2 materials did she perform as well as other students. In this instance, the PST would discuss the likelihood of any general education intervention being able to remediate a performance discrepancy of this magnitude, no matter how high intense or high quality. In this instance, a PST may forgo implementing and evaluating a general education intervention and proceed to an evaluation of special education entitlement.

Determining Educational Benefit From High-Quality Intervention

It is beyond the scope of this chapter to detail all the activities and conditions that must be in place to implement an RTI process with fidelity. (See Batsche et al., 2005, for general policies and procedures and Shinn, 2005a, for specific RTI processes that include CBM.) At the heart of the RTI process is the use of a scientifically based measure(s) that is sufficiently sensitive to improvement in achievement that judgments about educational benefit can be made in 4–10 weeks. This short time frame allows multiple interventions to be tested. CBM is well suited for this purpose.

See, for example, the progress-monitoring graph showing Georges' response to the selected general education intervention (Figure 10). The goal prescribes an expected rate of progress determined by his PST to be 2 WRC per week improvement. This expected rate of improvement is shown on the graph by a solid line.

Figure 8. Results of Georges' SLA on grades 1 and 2 R-CBM.

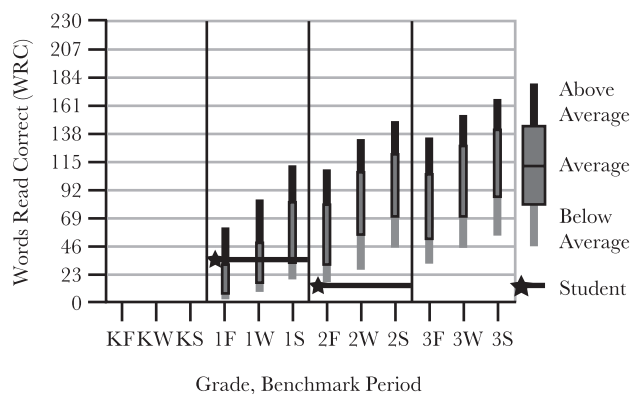


Figure 9. Results of Ginny's SLA showing a large performance discrepancy across peers and grades.

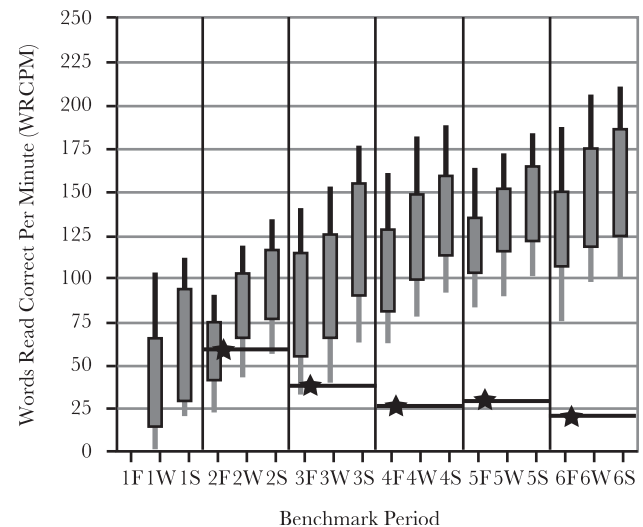
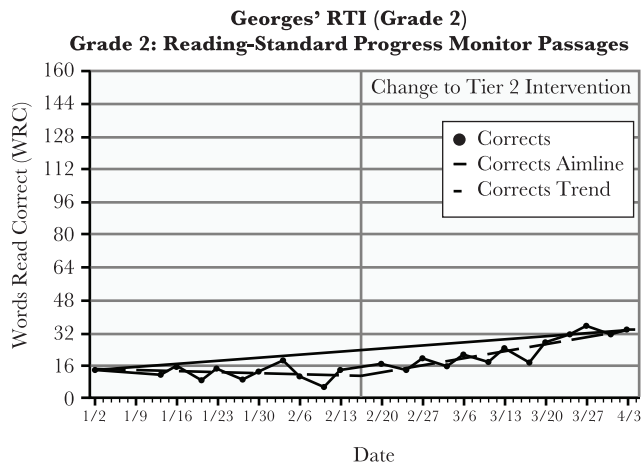


Figure 10. Results of Georges' RTI assessment.

With R-CBM, a student is tested one to two times per week with a single, but different, passage each time. Again, the trend line is computed using ordinary least squares (Good & Shinn, 1990) and is shown by the dashed line. For Georges, the actual rate of improvement was less than the expected rate of improvement when the effects of his current general education program were evaluated. Modification of his general education intervention included a high-quality Tier 2 intervention where he received an additional intensive supplemental program. Results of the change in intervention showed a dramatic increase in his reading performance, and he attained the goal that had been set for him as part of the RTI process. The PST would use these data to support a decision that the intervention was effective and that special education was not required for him to benefit. If Georges' response to a different intervention had not shown significant benefit, the intervention could be modified again or the need for special education intervention could be determined.

Exploring and Evaluating Solutions

If schools are well practiced in any assessment activity, it is in the area of testing for purposes of determining special education eligibility. Unfortunately, until recently, schools have been less proficient at assessing students' benefit from the services they receive as a result of the disability determination. Using CBM within a problem-solving model has contributed to a shift in assessment emphasis from just eligibility to one of improved outcomes (Ikeda et al., 2002; Tilly, Reschly, & Grimes, 1999). Using CBM in a problem-solving model allows for the generation of common data tied to

fundamental concepts. A student receives special education because of a severe educational need (i.e., significant performance discrepancy) and a failure to obtain educational benefit (i.e., rate of improvement) from high-quality general education intervention(s). These concepts are assessed first in general education and continue when a student receives special education. Fundamental to assessing educational benefit in this instance is progress toward the IEP annual goal.

As described in Shinn and Shinn (2000), most IEP goals have been written based on a mastery monitoring approach consisting of a variety of quasi-measurable short-term objectives (e.g., will master multiplication facts with 80% accuracy). The result has been what Fuchs and Fuchs (2002) detail as the production of IEPs that are onerous documents with an emphasis on procedure and process and that do not allow for accountability and progress. CBM IEP goals use a long-term goal approach to goal setting. In a long-term goal approach, measurement of progress toward the goal is designed to answer whether the student is becoming more proficient in reading, math, writing, or spelling in the general education curriculum. CBM goals replace a multitude of these short-term goals with a single outcome indicator in each academic area of deficit. Sample goals for reading or math computation would look like this: In 1 year, Ginny will read 120 WRC from grade 4 reading passages. In 1 year, Ginny will write 45 correct digits from grade 5 math computational problems.

A clear advantage of using CBM with students with disabilities to write IEP goals for students who receive special education is that it is a technology that has been *validated* for use in writing observable and measurable IEP goals and making statements about progress in general curriculum (Fuchs & Shinn, 1989). In fact, CBM's research-based development was supported by federal funds to identify a technology to assist in implementation of the 1975 Education for All Handicapped Children Act. For more detail on this history, see Deno (1992) or Shinn and Bamonto (1998).

The advantages of using CBM and a long-term goal approach are discussed in more detail in Fuchs and Fuchs (chapter 136, vol. 6) and in Fuchs (1993) and Fuchs and Deno (1992).

The goal-setting process requires specification of the (a) academic area in which the goal is to be written, (b) time frame in which the goal is expected to be accomplished, (c) level of performance at which the student will be expected to be proficient, and (d)

criterion for performance. The question for Ginny, the student shown in Figure 9, is how well do we want her to read in 1 year if her special education program is to be considered successful.

This goal-setting process begins by examining her performance on her SLA. Ginny is expected to be performing successfully on grade 6 reading passages. However, she is currently reading grade 2 passages as well as grade 2 students. In 1 year, the time frame for her IEP, her IEP team decided they wanted her to be successful reading grade 4 reading passages. The IEP team defined success as reading 90 WRC per minute on the grade 4 passages. This criterion would be equivalent to the reading skills of fourth-grade students.

The IEP team’s rationale for this goal was based on a discussion of reducing Ginny’s performance gap. In 1 year, she would be significantly closer to her peers. However, one disadvantage of this goal would be that she would still be likely to need an intensive intervention, albeit less likely. With CBM, a number of methods for this IEP goals setting process have been described in the professional literature (Fuchs & Shinn, 1989; Shinn & Shinn, 2000), including training manuals (Shinn, 2003a).

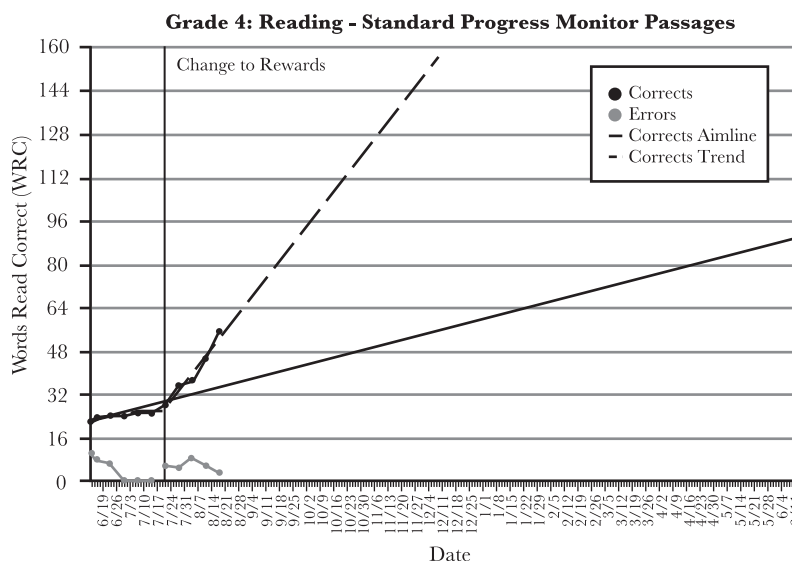
What is most important is that this goal translates into a picture of expected progress. The IEP annual goal, “in 1 year, Ginny will read 90 WRC in grade 4 reading passages” corresponds to the expected rate of progress (i.e., the aimline) on the student performance graph shown in Figure 11.

Progress-monitoring data usually are collected once or twice per week by a special education teacher or trained paraprofessional. In reading, students read one R-CBM probe from the goal material each testing session (i.e., grade 4). Ginny’s progress during the first intervention was below the expected rate of progress and, consistent with IDEA, her IEP was revised and noted on the graph. In this instance, Ginny’s reading intervention was supported by adding a component from the REWARDS intervention (Archer & Gleason, 2001). Progress monitoring after the intervention change showed a dramatic increase in reading skill. When the effects of instructional interventions are examined *systematically* and *continuously*, educators can make data-based decisions about whether to maintain or change interventions. Students need not receive instructional programs that do not meet their needs for long periods. Unsuccessful interventions can be changed to better meet students’ needs. Educators need not discard or modify programs that may be working. The net result is that student achievement outcomes are improved significantly (Fuchs et al., 1984; Fuchs & Fuchs, 1986, 1999).

Problem Solution

It can be argued persuasively that the kinds of testing activities that characterize long-term decisions about students’ educational needs (e.g., annual or 3-year reviews) are little more than expensive psychometric superstitious behavior. Annual reviews rely heavily on published achievements tests that were not designed to

Figure 11. Ginny’s progress toward her IEP annual goal, including revision to address lack of progress.

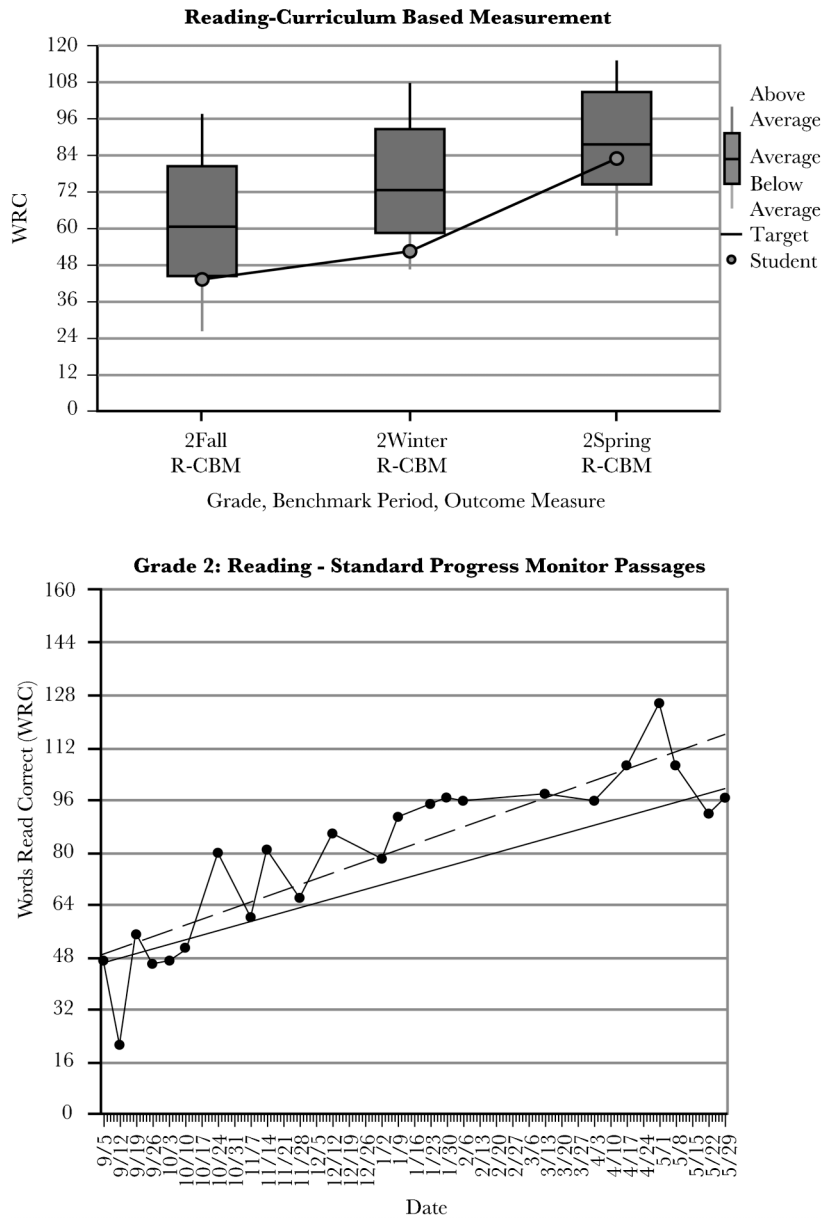


evaluate the progress of individual students (Carver, 1974; Marston, 1989). Special education exit decisions have occurred relatively infrequently, in large part, I believe, due to an inadequate system of progress monitoring. However, when CBM progress monitoring data are collected, it has been demonstrated that considerable numbers of students who received special education had attained academic skills equal to those of general education students. For example, in at least three school districts, 30–40% of special education students served more than 2 years in a pull-out program with a reading IEP objective read as well or better than

other low-performing readers in general education classroom (Powell-Smith & Habedank, 1998; Shinn, Baker, Habedank, & Good, 1993; Shinn, Habedank, Rodden-Nord, & Knutson, 1993). These students are appropriate candidates for exit from special education. (See Powell-Smith & Ball, chapter 15, vol. 2, for more information.)

A problem-solving model resolves many problems in determining students' long-term needs and the appropriateness of continued special education services. The severity of educational need and lack of educational benefit that defined the need for special education

Figure 12. Annual review showing significantly decreased educational need (reduced discrepancy) and educational benefit (progress toward IEP goal).



remain the basis for decision making. The questions at periodic, annual, and 3-year reviews focus on the following:

- Is there educational benefit? Is the student making the expected rate of progress toward the IEP goal? If yes, should the IEP goal be raised? If not, what changes in intervention are required?
- Is there still severe educational need? Is the student benefiting from the intervention such that the student is reducing the performance discrepancy? If yes, does the student still need special education to benefit? If no, what changes in intervention are required?

At periodic and annual reviews, benchmark data and progress toward IEP goals are examined systematically as illustrated in Figure 12. The first graph shows John's grade 2 fall, winter, and spring benchmark scores. The second graph shows his rate of progress toward his reading IEP goal. Because of severe educational need, he had been placed in special education in kindergarten but by the beginning of fall of second grade had significantly reduced his performance discrepancy.

By the end of grade 2, at the annual review, the performance discrepancy had been eliminated as he read within the range of his general education peers. His rate of progress toward his IEP goal also exceeded the expected rate of progress. Two sources of data indicated significant educational benefit, and the benchmark results documented that John no longer needed special education. (See Powell-Smith & Habedank, 1998; Shinn, Powell-Smith, Good, & Baker, 1997; and Shinn, Powell-Smith, & Good, 1996)

SUMMARY

The times are changing. In contrast to previous years where CBM appealed to a sizable number, although still a minority, of school psychologists, changes in knowledge and changes in educational law and policies have made it a standard tool in data-based decision making for all students in the basic skill areas. These changes have made CBM of interest to all school psychologists to contribute to decisions about educational need and educational benefits on a continuous basis. This ability to attend to educational benefit in particular allows school psychologists to expand their role to support personnel and interventions to help solve problems rather than just identify or admire them.

REFERENCES

- Archer, A., & Gleason, M. (2001). *REWARDS: Reading Excellence Word Attack and Rate Development Strategies*. Longmont, CO: Sopris West.
- Bateman, B. D., & Linden, M. A. (1998). *Better IEPs*. Longmont, CO: Sopris West.
- Batsche, G. M., Elliott, J., Graden, J., Grimes, J., Kovaleski, J. F., Prasse, D., et al. (2005). *Response to intervention: Policy considerations and implementation*. Alexandria, VA: National Association of State Directors of Special Education.
- Carlberg, C., & Kavale, K. A. (1980). The efficacy of special versus regular class placement for exceptional children: A meta-analysis. *Exceptional Children, 14*, 295–308.
- Carver, R. P. (1974). Two dimensions of tests: Psychometric and edumetric. *American Psychologist, 29*, 512–518.
- Deno, D., Deno, S. L., Marston, D., & Marston, D. B. (1987). *Test of Oral Reading Fluency*. Eden Prairie: Children's Educational Services.
- Deno, S. L. (1985). Curriculum-based measurement: The emerging alternative. *Exceptional Children, 52*, 219–232.
- Deno, S. L. (1986). Formative evaluation of individual student programs: A new role for school psychologists. *School Psychology Review, 15*, 358–374.
- Deno, S. L. (1989). Curriculum-based measurement and alternative special education services: A fundamental and direct relationship. In M. R. Shinn (Ed.), *Curriculum-based measurement: Assessing special children* (pp. 1–17). New York: Guilford Press.
- Deno, S. L. (1992). The nature and development of curriculum-based measurement. *Preventing School Failure, 36*(2), 5–10.
- Deno, S. L. (1995). School psychologist as problem solver. In A. Thomas & J. Grimes (Eds.), *Best practices in school psychology III* (pp. 471–484). Washington, DC: National Association of School Psychologists.
- Deno, S. L. (2002). Problem-solving as best practice. In A. Thomas & J. Grimes (Eds.), *Best practices in school psychology IV* (pp. 37–55). Bethesda, MD: National Association of School Psychologists.
- Deno, S. L. (2003). Developments in curriculum-based measurement. *The Journal of Special Education, 37*, 184–192.
- Deno, S. L. (2005). Problem-solving assessment. In R. Brown-Chidsey (Ed.), *Assessment for intervention: A problem-solving approach* (pp. 10–40). New York: Guilford Press.
- Deno, S. L., Fuchs, L. S., Marston, D., & Shin, J. (2001). Using curriculum-based measurement to establish growth standards for students with learning disabilities. *School Psychology Review, 30*, 507–524.
- Deno, S. L., Marston, D., & Mirkin, P. (1982). Valid measurement procedures for continuous evaluation of written expression. *Exceptional Children, 48*, 68–71.

- Deno, S. L., & Mirkin, P. (1980). Data-based IEP development: An approach to substantive compliance. *Teaching Exceptional Children, 12*, 92–97.
- Deno, S. L., Mirkin, P., & Chiang, B. (1982). Identifying valid measures of reading. *Exceptional Children, 49*, 36–45.
- Foegen, A., Jiban, C., & Deno, S. L. (2007). Progress monitoring measures in mathematics: A review of the literature. *The Journal of Special Education, 41*, 121–139.
- Fuchs, D., Fuchs, L. S., Mathes, P. G., Lipsey, M. E., & Eaton, S. (2000). A meta-analysis of reading differences between underachievers with and without the learning disabilities label: A brief report. *Learning Disabilities: A Multidisciplinary Journal, 10*, 1–3.
- Fuchs, D., Fuchs, L. S., Mathes, P. G., Lipsey, M. W., & Roberts, P. H. (2001, August). Is “learning disabilities” just a fancy term for low achievement? A meta-analysis of reading differences between low achievers with and without the label: Executive summary. Paper presented at the Learning Disabilities Summit: Building a Foundation for the Future, Washington, DC.
- Fuchs, L. S. (1986). Monitoring progress among mildly handicapped pupils: Review of current practice and research. *Remedial and Special Education, 7*, 5–12.
- Fuchs, L. S. (1993). Enhancing instructional programming and student achievement with curriculum-based measurement. In J. Kramer (Ed.), *Curriculum-based measurement* (pp. 65–104). Lincoln, NE: Buros Institute of Mental Measurements.
- Fuchs, L. S. (1998). Computer applications to address implementation difficulties associated with curriculum-based measurement. In M. R. Shinn (Ed.), *Advanced applications of curriculum-based measurement* (pp. 89–112). New York: Guilford Press.
- Fuchs, L. S., & Deno, S. L. (1992). Effects of curriculum within curriculum-based measurement. *Exceptional Children, 58*, 232–243.
- Fuchs, L. S., & Deno, S. L. (1994). Must instructionally useful performance assessment be based in the curriculum? *Exceptional Children, 61*, 15–24.
- Fuchs, L. S., Deno, S. L., & Mirkin, P. (1984). The effects of frequent curriculum-based measurement and evaluation on pedagogy, student achievement and student awareness of learning. *American Educational Research Journal, 21*, 449–460.
- Fuchs, L. S., & Fuchs, D. (1986). Effects of systematic formative evaluation: A meta-analysis. *Exceptional Children, 53*, 199–208.
- Fuchs, L. S., & Fuchs, D. (1988). Curriculum-based measurement: A methodology for evaluating and improving student programs. *Diagnostic, 14*, 3–13.
- Fuchs, L. S., & Fuchs, D. (1992). Identifying a measure for monitoring student reading progress. *School Psychology Review, 21*, 45–58.
- Fuchs, L. S., & Fuchs, D. (1999). Monitoring student progress toward the development of reading competence: A review of three forms of classroom-based assessment. *School Psychology Review, 28*, 659–671.
- Fuchs, L. S., & Fuchs, D. (2002). *Progress monitoring, accountability, and LD identification*. Testimony to the President’s Commission on Excellence in Special Education, Washington, DC.
- Fuchs, L. S., & Fuchs, D. (2004). *What is scientifically based research on progress monitoring?* Washington, DC: National Center on Progress Monitoring, American Institute for Research, Office of Special Education Programs.
- Fuchs, L. S., Fuchs, D., & Deno, S. L. (1982). Reliability and validity of curriculum-based informal reading inventories. *Reading Research Quarterly, 18*, 6–26.
- Fuchs, L. S., Fuchs, D., & Hamlett, C. L. (1989). Effects of instrumental use of curriculum-based measurement to enhance instructional programs. *Remedial and Special Education, 10*, 43–52.
- Fuchs, L. S., Fuchs, D., & Hamlett, C. L. (1995). *Monitoring basic skills progress (MBSP): Basic math computation and basic math concepts and applications*. Austin, TX: PRO-ED.
- Fuchs, L. S., Fuchs, D., Hamlett, C. L., & Allinder, R. M. (1991). The contribution of skills analysis to curriculum-based measurement in spelling. *Exceptional Children, 57*, 443–452.
- Fuchs, L. S., Fuchs, D., Hamlett, C. L., & Ferguson, C. (1992). Effects of expert system consultation within curriculum-based measurement using a reading maze task. *Exceptional Children, 58*, 436–450.
- Fuchs, L. S., Fuchs, D., Hamlett, C. L., & Stecker, P. M. (1991). Effects of curriculum-based measurement and consultation on teaching planning and student achievement in mathematics operations. *American Educational Research Journal, 28*, 617–641.
- Fuchs, L. S., Fuchs, D., Hamlett, C. L., Thompson, A., Roberts, P. H., & Kupek, P. (1994). Technical features of a mathematics concepts and applications curriculum-based measurement system. *Diagnostic, 19*, 23–49.
- Fuchs, L. S., Fuchs, D., Hosp, M. K., & Jenkins, J. R. (2001). Oral Reading Fluency as an indicator of reading competence: A theoretical, empirical, and historical analysis. *Scientific Studies of Reading, 5*, 239–256.
- Fuchs, L. S., Fuchs, D., & Maxwell, L. (1988). The validity of informal reading comprehension measures. *Remedial and Special Education, 9*, 20–28.
- Fuchs, L. S., Fuchs, D., & Speece, D. L. (2002). Treatment validity as a unifying construct for identifying learning disabilities. *Learning Disability Quarterly, 25*, 33–45.
- Fuchs, L. S., & Shinn, M. R. (1989). Writing CBM IEP objectives. In M. R. Shinn (Ed.), *Curriculum-based measurement: Assessing special children* (pp. 132–154). New York: Guilford Press.
- Germann, G., & Tindal, G. (1985). An application of curriculum based assessment: The use of direct and repeated measurement. *Exceptional Children, 52*, 244–265.

- Good, R. H., III, & Jefferson, G. (1998). Contemporary perspectives on curriculum-based measurement validity. In M. R. Shinn (Ed.), *Advanced applications of curriculum-based measurement* (pp. 61–88). New York: Guilford Press.
- Good, R. H., III, & Shinn, M. R. (1990). Forecasting accuracy of slope estimates for reading curriculum-based measurement: Empirical evidence. *Behavioral Assessment, 12*, 179–193.
- Gottlieb, J., Alter, M., Gottlieb, B. W., & Wishner, J. (1994). Special education in urban America. It's not justifiable for many. *The Journal of Special Education, 27*, 453–465.
- Gresham, F. M., MacMillan, D. L., & Bocian, K. M. (1996). Learning disabilities, low achievement, and mild mental retardation: More alike than different? *Journal of Learning Disabilities, 29*, 570–581.
- Hamilton, C., & Shinn, M. R. (2003). Characteristics of word callers: An investigation of the accuracy of teachers' judgments of reading comprehension and oral reading skills. *School Psychology Review, 32*, 228–240.
- Hintze, J. M., Shapiro, E. S., & Daly, E. J. (1998). An investigation of the effects of passage difficulty level on outcomes of Oral Reading Fluency progress monitoring. *School Psychology Review, 27*, 433–445.
- Hintze, J. M., Shapiro, E. S., & Lutz, J. G. (1994). The effects of curriculum on the sensitivity of curriculum-based measurement of reading. *The Journal of Special Education, 28*, 188–202.
- Hintze, J. M., & Silbergitt, B. (2005). A longitudinal examination of the diagnostic accuracy and predictive validity of R-CBM and high stakes testing. *School Psychology Review, 34*, 372–386.
- Hosp, M. K. (2005). Using CBM as an indicator of decoding, word reading, and comprehension: Do the relations change with grade? *School Psychology Review, 34*, 9–26.
- Howe, K. B., & Shinn, M. M. (2002). *Standard reading assessment passages (RAPs) for use in general outcome measurement: A manual describing development and technical features*. Eden Prairie, MN: Edformation.
- Howell, K. W., Kurns, S., & Antil, L. (2002). Best practices in curriculum-based evaluation. In A. Thomas & J. Grimes (Eds.), *Best practices in school psychology IV* (pp. 671–698). Bethesda, MD: National Association of School Psychologists.
- Howell, K. W., & Nolet, V. (1999). *Curriculum-based evaluation: Teaching and decision making* (3rd ed.). Atlanta, GA: Wadsworth.
- Ikeda, M. J., Grimes, J., Tilly, W. D., III, Allison, R., Kurns, S., Stumme, J., et al. (2002). Implementing an intervention-based approach to service delivery: A case example. In M. R. Shinn, H. M. Walker, & G. Stoner (Eds.), *Interventions for academic and behavior problems: Preventive and remedial approaches* (pp. 71–88). Bethesda, MD: National Association of School Psychologists.
- Ikeda, M. J., Tilly, W. D., III, Stumme, D., Volmer, L., & Allison, R. (1996). Agency-wide implementation of problem-solving consultation: Foundations, current implementation, and future directions. *School Psychology Quarterly, 11*, 228–243.
- Jenkins, J. R., Deno, S. L., & Mirkin, P. K. (1979). Measuring pupil progress toward the least restrictive environment. *Learning Disability Quarterly, 2*, 81–92.
- Kaminski, R. A., & Good, R. H., III. (1996). Toward a technology of assessing basic early literacy skills. *School Psychology Review, 25*, 215–227.
- Kaminski, R. A., & Good, R. H., III. (1998). Assessing early literacy skills in a problem-solving model: Dynamic Indicators of Basic Skills. In M. R. Shinn (Ed.), *Advanced applications of curriculum-based measurement* (pp. 113–142). New York: Guilford Press.
- Kavale, K., & Forness, S. R. (1985). *The science of learning disabilities*. San Diego, CA: College-Hill Press.
- Kavale, K. A., & Forness, S. R. (1999). Effectiveness of special education. In C. R. Reynolds & T. B. Gutkin (Eds.), *The handbook of school psychology* (3rd ed., pp. 984–1024). New York: John Wiley.
- Marston, D. (1989). Curriculum-based measurement: What is it and why do it? In M. R. Shinn (Ed.), *Curriculum-based measurement: Assessing special children* (pp. 18–78). New York: Guilford Press.
- Marston, D., & Magnusson, D. (1985). Implementing curriculum-based measurement in special and regular education settings. *Exceptional Children, 52*, 266–276.
- McMaster, K., & Espin, C. (2007). Technical features of curriculum-based measurement in writing. *The Journal of Special Education, 41*, 68–84.
- Miura Wayman, M., Wallace, T., Ives Wiley, H., Ticha, R., & Espin, C. (2007). Literature synthesis on curriculum-based measurement in reading. *The Journal of Special Education, 41*, 85–120.
- Pericola Case, L., Speece, D. L., & Eddy Molloy, D. (2003). The validity of a response-to-instruction paradigm to identify reading disabilities: A longitudinal analysis of individual differences and context factors. *School Psychology Review, 32*, 557–582.
- Peterson, K. M., & Shinn, M. R. (2002). Severe discrepancy models: Which best explains school identification practices for learning disabilities? *School Psychology Review, 31*, 459–476.
- Powell-Smith, K., & Habedank, L. (1998). The use of curriculum-based measurement in the reintegration of students with mild disabilities. In M. R. Shinn (Ed.), *Advanced applications of curriculum-based measurement* (pp. 254–307). New York: Guilford Press.
- Powell-Smith, K., & Shinn, M. R. (2004). *AIMSweb training workbook administration and scoring of written expression curriculum-based measurement (WE-CBM) for use in general outcome measurement*. Eden Prairie, MN: Edformation.
- Reynolds, M. C., & Heistad, D. (1997). 20/20 analysis: Estimating school effectiveness in serving students at the margins. *Exceptional Children, 63*, 439–449.
- Shapiro, E. S. (2004). *Academic skills problems: Direct assessment and intervention* (3rd ed.). New York: Guilford Press.

- Shinn, M. R. (Ed.). (1989). *Curriculum-based measurement: Assessing special children*. New York: Guilford Press.
- Shinn, M. R. (1995). Curriculum-based measurement and its use in a problem-solving model. In A. Thomas & J. Grimes (Eds.), *Best practices in school psychology III*. Washington, DC: National Association of School Psychologists.
- Shinn, M. R. (Ed.). (1998). *Advanced applications of curriculum-based measurement*. New York: Guilford Press.
- Shinn, M. R. (2002). Best practices in curriculum-based measurement and its use in a problem-solving model. In A. Thomas & J. Grimes (Eds.), *Best practices in school psychology IV* (pp. 671–698). Bethesda, MD: National Association of School Psychologists.
- Shinn, M. R. (2003a). *AIMSweb training workbook progress-monitoring strategies for writing individualized goals in general curriculum and more frequent formative evaluation*. Eden Prairie, MN: Edformation.
- Shinn, M. R. (2003b). *AIMSweb training workbook: Organizing and implementing a benchmark assessment program*. Eden Prairie, MN: Edformation.
- Shinn, M. R. (2005a). *AIMSweb response to intervention (RTI): A standard protocol-based system for managing and reporting problem-solving outcomes*. Eden Prairie, MN: Edformation.
- Shinn, M. R. (2005b). Identifying and validating academic problems. In R. Brown-Chidsey (Ed.), *Assessment for intervention: A problem-solving approach* (pp. 219–246). New York: Guilford Press.
- Shinn, M. R., Baker, S., Habedank, L., & Good, R. H., III. (1993). The effects of classroom reading performance data on general education teachers' and parents' attitudes about reintegration. *Exceptionality, 4*, 205–228.
- Shinn, M. R., & Bamonto, S. (1998). Advanced applications of curriculum-based measurement: “Big ideas” and avoiding confusion. In M. R. Shinn (Ed.), *Advanced applications of curriculum-based measurement* (pp. 1–31). New York: Guilford Press.
- Shinn, M. R., Good, R. H., III, Knutson, N., Tilly, W. D., III, & Collins, V. (1992). Curriculum-based reading fluency: A confirmatory analysis of its relation to reading. *School Psychology Review, 21*, 458–478.
- Shinn, M. R., Good, R. H., III, & Parker, C. (1999). Non-categorical special education services with students with severe achievement deficits. In D. J. Reschly, W. D. Tilly III, & J. P. Grimes (Eds.), *Special education in transition: Functional assessment and noncategorical programming* (pp. 81–106). Longmont, CO: Sopris West.
- Shinn, M. R., Good, R. H., III, & Stein, S. (1989). Summarizing trend in student achievement: A comparison of evaluative models. *School Psychology Review, 18*, 356–370.
- Shinn, M. R., Habedank, L., Rodden-Nord, K., & Knutson, N. (1993). Using curriculum-based measurement to identify potential candidates for reintegration into general education. *The Journal of Special Education, 27*, 202–221.
- Shinn, M. R., Powell-Smith, K. A., & Good, R. H., III. (1996). Evaluating the effects of responsible reintegration into general education for students with mild disabilities on a case-by-case basis. *School Psychology Review, 25*, 519–539.
- Shinn, M. R., Powell-Smith, K. A., Good, R. H., III, & Baker, S. (1997). The effects of reintegration into general education reading instruction for students with mild disabilities. *Exceptional Children, 64*, 59–79.
- Shinn, M. R., Rosenfield, S., & Knutson, N. (1989). Curriculum-based assessment: A comparison and integration of models. *School Psychology Review, 18*, 299–316.
- Shinn, M. R., & Shinn, M. M. (2000). Writing and evaluating IEP goals and making appropriate revisions to ensure participation and progress in general curriculum. In C. F. Tetzlow & M. Tankersley (Eds.), *IDEA amendments of 1997: Practice guidelines for school-based teams* (pp. 351–381). Bethesda, MD: National Association of School Psychologists.
- Shinn, M. R., & Shinn, M. M. (2003a). *AIMSweb training workbook administration and scoring of reading curriculum-based measurement (R-CBM) for use in general outcome measurement*. Eden Prairie, MN: Edformation.
- Shinn, M. R., & Shinn, M. M. (2003b). *AIMSweb training workbook administration and scoring of reading maze for use in general outcome measurement*. Eden Prairie, MN: Edformation.
- Shinn, M. R., & Shinn, M. M. (2003c). *AIMSweb training workbook administration and scoring of spelling curriculum-based measurement (S-CBM) for use in general outcome measurement*. Eden Prairie, MN: Edformation.
- Shinn, M. R., & Shinn, M. M. (2004). *AIMSweb training workbook administration and scoring of mathematics curriculum-based measurement (M-CBM) for use in general outcome measurement*. Eden Prairie, MN: Edformation.
- Shinn, M. R., Shinn, M. M., Hamilton, C., & Clarke, B. (2002). Using curriculum-based measurement to promote achievement in general education classrooms. In M. R. Shinn, G. Stoner, & H. M. Walker (Eds.), *Interventions for academic and behavior problems: Preventive and remedial approaches* (pp. 113–142). Bethesda, MD: National Association of School Psychologists.
- Shinn, M. R., Tindal, G., Spira, D., & Marston, D. (1987). Practice of learning disabilities as social policy. *Learning Disability Quarterly, 10*, 17–28.
- Shinn, M. R., Ysseldyke, J., Deno, S. L., & Tindal, G. (1986). A comparison of differences between students labeled learning disabled and low achieving on measures of classroom performance. *Journal of Learning Disabilities, 19*, 545–552.
- Silberglitt, B., & Hintze, J. M. (2005). Formative assessment using CBM-R cut score to track progress toward success on state-mandated achievement tests: A comparison of methods. *Journal of Psychoeducational Assessment, 23*, 304–325.

- Smith, S. W. (1990). Individualized educational programs (IEPs) in special education: From intent to acquiescence. *Exceptional Children*, 57, 6–14.
- Tilly, W. D., III, Reschly, D. J., & Grimes, J. (1999). Disability determination in problem-solving systems: Conceptual foundations and critical components. In D. J. Reschly, W. D. Tilly III, & J. P. Grimes (Eds.), *Special education in transition: Functional assessment and noncategorical programming* (pp. 221–254). Longmont, CO: Sopris West.
- Tindal, G. (1993). A review of curriculum-based procedures on nine assessment components. In J. Kramer (Ed.), *Curriculum-based measurement* (pp. 25–64). Lincoln, NE: Buros Institute of Mental Measurements.
- Tindal, G., Wesson, C., Deno, S. L., Germann, G., & Mirkin, P. (1985). The Pine County model for special education delivery: A data-based system. In T. Kratochwill (Ed.), *Advances in school psychology* (Vol. 4, pp. 223–250). Hillsdale, NJ: Erlbaum.
- Ysseldyke, J., Burns, M., Dawson, P., Kelley, B., Morrison, D., Ortiz, S., et al. (2006). *School psychology: A blueprint for training and practice III*. Bethesda, MD: National Association of School Psychologists.
- Fuchs, L. S. (1994). *Connecting performance assessment to instruction*. Reston, VA: Council for Exceptional Children.
- Lays out the basic details of the necessary features of any assessment system that is designed to improve student achievement outcomes. Should be standard reading for anyone interested in assessing student achievement for instructional planning and educational outcomes.
- Fuchs, L. S., & Deno, S. L. (1991). Paradigmatic distinctions between instructionally relevant measurement models. *Exceptional Children*, 58, 232–243.
- Serves as a primer for comparing the advantages and disadvantages of types of outcome monitoring assessment systems, including specific subskill mastery measurement (short-term goal approach) and general outcome measurement (long-term goal approach) progress monitoring. General outcome measurement emerges as the most advantageous method for monitoring student achievement outcomes according to a set of standards that includes technical adequacy and feasibility.
- Shinn, M. R. (Ed.). (1998). *Advanced applications of curriculum-based measurement*. New York: Guilford Press.
- Presents information on a range of applications of CBM to problem solving, including points of confusion, use with English language learners and computers.
- Shinn, M. R. (Ed.). (1989). *Curriculum-based measurement: Assessing special children*. New York: Guilford Press.
- Contains contributions of the work of the core group of researchers and school practitioners (Deno, L. Fuchs, Marston, Shinn, Tindal, Allen) who developed and implemented CBM. Specific procedures for implementing CBM in school settings are detailed. General background information and conceptual issues surrounding the development and use of CBM as an alternative to traditional school psychology practice are discussed. A case study is presented to illustrate the differences between approaches.

ANNOTATED BIBLIOGRAPHY

- Deno, S. L. (1991). Individual differences and individual difference: The essential difference of special education. *The Journal of Special Education*, 24, 160–173.

The current aptitude–treatment interaction instructional technology is examined and determined to be insufficiently responsive to individual differences of students. Instead, a formative evaluation model is proposed as an idiographic approach to building more effective programs for individual students. In this model, aptitude differences at the outset of a program are deemphasized, and methodology for tailoring programs to individual students during instruction becomes the focus.