CHAPTER 11

Building a Scientifically Based Data System for Progress Monitoring and Universal Screening Across Three Tiers, Including RTI Using Curriculum-Based Measurement

Mark R. Shinn
National-Louis University

INTRODUCTION

In the past decade, educational science and quality school-based practices have come together to develop more unified, positive, and preventive service delivery systems for academic performance and behavioral support in America's schools. Among the numerous labels for these service delivery systems are early-intervening services, coordinated early-intervening services (CEIS), multitiered early-intervening services, and a three-tier problem-solving model. However, the most common term for these types of service delivery systems is *response to intervention* (RTI).

The term RTI has generated considerable controversy and confusion. To some in the field, the term describes an allowable, if not encouraged, entitlement process to determine special education eligibility for specific learning disabilities (SLD). For many others, RTI is a generic term for a much broader service delivery system that has among its components (a) multitiered interventions (e.g., three tiers) of increasing intensity to address academic performance and behavior, and (b) data-based decision making, including universal screening and progress monitoring. This latter use of RTI constitutes the context for this chapter. The chapter also illustrates how curriculum-based measurement (CBM), a set of simple, time-efficient, and scientifically based basic skills tests, can be used across multiple tiers for universal screening and progress monitoring. The academic area of reading is used as primary example.

Interventions INT_Ch10_11.3d 2/2/10 19:55:34
The Charlesworth Group, Wakefield +44(0)1924 369598 - Rev 7.51n/W (Jan 20 2003)

OVERVIEW OF UNIVERSAL SCREENING AND PROGRESS MONITORING

Over the past 30 years, a large body of knowledge has been generated that demonstrates what methods enable students to become competent readers. Adams (1990) noted that to be able to read well, children must read widely, and in order for them to read widely, they must be able to read well. Furthermore, good and wide reading must be accomplished early. Torgesen (1998, 2004) has reported that in order to ensure this healthy start, schools must (a) increase the quality, consistency, and ''reach'' of instruction in every kindergarten–Grade 3 classroom; (b) engage in *universal screening* and timely and valid assessments of reading growth (i.e., *frequent progress monitoring*); and (c) provide more intensive interventions to ''catch up'' struggling readers.

This chapter focuses on the second requirement, while chapters 18 and 6 in this book devote considerable attention to the first and third. Not only are universal screening and progress monitoring important for getting all students off to a healthy start in reading in Grades K–3; they also are foundational components in multitiered coordinated early-intervening services. In addition, progress monitoring is both a legal requirement and ''best practice'' in RTI as a special education entitlement process for specific learning disabilities (Burdette, 2007; Fuchs & Vaughn, 2005).

Common Universal Screening Practices in a Multitiered CEIS Model

In universal screening, *all* students are tested using a standard measure for determining which learners may be sufficiently different from expectations to warrant more intensive interventions. Universal screening differs from individual screening, which involves testing a subset of students, usually one at a time, in a process that typically is initiated by teacher referral. Although teacher referral is generally accurate for identifying academic problems (Algozzine, Christenson, & Ysseldyke, 1982; Gerber & Semmel, 1984) and behavior problems (Walker & Severson, 1994), it has been shown that this process has bias for both gender and ethnicity (Shinn, Tindal, & Spira, 1987). By testing all students and eliminating identification by referral, individual teacher bias can be minimized.

Schools, of course, have considerable experience with universal screening in areas other than academics. For example, universal hearing and vision screening have been common educational practice for more than 50 years. When schools engage in the process of implementing multitier service delivery systems, they go beyond hearing and vision screening and expand universal screening to reading. When they do so, two universal screening approaches for reading might be employed that can be likened to the medical treatment approaches (a) *titration* and (b) *triage*. In both approaches, decisions to provide more intensive interventions are based on set criterion, or ''cut scores.'' Most often, these cut scores are *normative*, and students performing below a certain percentile (e.g., the 25th percentile) receive more intensive intervention. In

Interventions INT_Ch10_11.3d 2/2/10 19:55:35
The Charlesworth Group, Wakefield +44(0)1924 369598 - Rev 7.51n/W (Jan 20 2003)

other instances, the cut score is based on level of performance tied to a *standard*. For example, a standards-based cut score is often used with Dynamic Indicators of Basic Early Literacy (DIBELS; Kaminski & Good, 1998), where on a measure such as phonemic segmentation fluency (PSF), performance below a specified value (e.g., 35) would suggest the need for Tier 2 intervention.

Universal Screening as Titration

According to Wikipedia, *titration* is a medical term describing the process of ''gradually adjusting the dose of a medication until the desired effect is achieved.'' In this most frequently occurring universal screening approach, all students are tested, and those below the cut score are identified as potential candidates for more intensive Tier 2 reading intervention. This approach is called titration because Tier 3 interventions are not provided until students have been shown not to respond to Tier 2 interventions (Fletcher, Coulter, Reschly, & Vaughn, 2004; Fuchs & Fuchs, 2005; Vaughn & Fuchs, 2003). In other words, the *dosage*—the intensity of treatment—is based on a judgment regarding whether the Tier 2 intervention had its desired effect. Furthermore, the decision to change tiered interventions is not based on the universal screening data, but on progress monitoring data. In an extension of the titration metaphor, those students who do not respond to Tier 3 may be considered for an even more intensive dosage, that is, special education entitlement, as part of RTI. Although this titration model is defensible and quite common, it can be viewed as another ''wait-to-fail'' approach, albeit based on student achievement data rather than on an ability–achievement discrepancy.

Universal Screening as Triage

The second, less common approach to universal screening is referred to as *triage*. According to Wikipedia, triage is a process used in a ''scene of mass casualty, in order to sort patients into those who need critical attention and immediate transport to the hospital and those with less serious injuries.'' When universal screening data in reading are used in triage, students' scores allow for immediate alignment with *any* of the tiers appropriate to the severity of their needs. In a three-tier model, students whose scores are average or above receive the core reading instructional program (i.e., Tier 1). Students whose scores are below average (e.g., 25th percentile) may receive Tier 2 reading intervention in addition to their Tier 1 program. Students with the greatest reading needs (e.g., below the 10th percentile) may receive the most powerful, intense intervention, Tier 3. In a triage approach, there is no need to fail at a particular tier before receiving a more intensive intervention. Instead, students are provided services at an appropriate level as soon as the need is identified.

Common and Preferred Universal Screening Tools

Within certain parameters (e.g., sensitivity, specificity), nearly any reliable and valid achievement test *may* be suitable for use in universal screening. Typically, schools select

a reading screening tool that is (a) already in use in general education classrooms, and (b) consistent with a district's reading philosophy. Examples of measures that are employed include high-stakes state tests, extant group-administered achievement tests, individually administered achievement tests (both lengthy and short), computerized tests, and a variety of informal assessment strategies such as running records. Unfortunately, schools too rarely consider time efficiency and cost-effectiveness when selecting their screening tools.

For best practices, schools instead should first consider the potential screener's technical adequacy, and then include time and cost-efficient considerations. Ideally, students with more intensive reading needs would be identified accurately with minimal loss of instructional time and financial expense. Significant value to the screener can be added if the tool is also validated for progress monitoring.

Common Progress Monitoring Practices in a Multitiered Model

In practice, less attention has been directed toward schools' systems of progress monitoring in reading, despite what one would expect. One would assume that schools are adept at monitoring academic progress, given that the purpose of schools is learning. However, it has been well established that monitoring the progress of learners has been, and remains, an area of weakness in instructional practice (Baker, O'Neil, & Linn, 1993; Deno, 1986, 2005; Fuchs & Fuchs, 1984, 2008).

When schools begin to employ progress monitoring systems across multiple tiers, three broad approaches have been identified, referred to here as (a) unspecified and discontinuous progress monitoring, or ''independent contracting''; (b) specified, continuous, but inadequate progress monitoring; and (c) specified, continuous, and adequate progress monitoring.

Independent Contracting Progress Monitoring

In spite of the emphasis given by the No Child Left Behind Act of 2001 (NCLB) and the Individuals with Disabilities Education Act (IDEA 2004) to assessing student achievement over time, some schools continue to pay little attention to progress monitoring. These progress monitoring practices can be described as unspecified because they lack an overall plan for implementation. In these schools, any progress monitoring practices that may exist are ones that have been in use for years. Progress monitoring also is discontinuous, because each intervention program or tier uses different progress monitoring tools. It may seem as though every program is conducting monitoring its own way, or what could be called ''independent contracting.''

For example, general education (Tier 1) may use group achievement tests, curriculum-embedded assessments, informal tests, or repeated administrations (e.g., three to four) of the universal screening test. In Tier 2, different progress monitoring approaches may be used, such as curriculum-embedded assessments from a different

Interventions INT_Ch10_11.3d 2/2/10 19:55:35
The Charlesworth Group, Wakefield +44(0)1924 369598 - Rev 7.51n/W (Jan 20 2003)

curriculum or informal teacher-made tests. In Tier 3, progress monitoring approaches most commonly include lengthy and infrequent (e.g., once per year) individualized achievement tests and/or yet another type of curriculum-embedded assessments.

In this independent contracting approach, there is little effort to build a *comprehensive*, systematic service delivery system with the features of coordinated early-intervening services (Burdette, 2007). There is little emphasis on processes to evaluate systematically if the solution was effective or needed modification. Furthermore, many of the traditional progress monitoring methods that are employed are not scientifically based (Fuchs & Fuchs, 2008); as a result, decisions about when or if a student has responded to an intervention (i.e., made adequate progress) may be contentious and driven by opinion.

Specified, Continuous, but Inadequate Progress Monitoring

In settings where there has been an effort to specify progress monitoring tools and to use the same tests and data base across intervention tiers or programs (i.e., building *continuous* progress monitoring practices), it is common for schools to build their progress monitoring process for reading using their universal screening tests. The presumption is that if the universal screening reading test is administered over time, it is valid for making progress decisions in general education and at different tiers. This approach is more prescriptive and has more continuity than the independent contracting approach, because all programs or tiers use the same data to evaluate progress. However, too often, the tests used may not be technically adequate (e.g., reliable, valid) for progress monitoring purposes. Because IDEA 2004 requires that assessments be technically adequate, legal and practice problems may arise when measures that lack these features are used to determine (a) progress toward individual education plan (IEP) goals (Fuchs, Fuchs, & Deno, 1985; Shinn & Shinn, 2000) and (b) response to intervention as part of SLD eligibility determination (Fuchs & Vaughn, 2005; Pericola Case, Speece, & Eddy Molloy, 2003; Vaughn & Fuchs, 2003).

The primary difficulty in creating a progress monitoring system based on universal screening reading tests, such as high-stakes state tests or group achievement tests, is that these measures are *summative* tests. That is, they were designed to assess the effects of instruction *after* instruction has occurred. Summative tests are constructed to be sensitive to *between*-person differences (Howell, Kurns, & Antil, 2002; Howell & Nolet, 1999). In other words, a valid summative reading test that is useful for universal screening should distinguish between poor readers and good readers. Summative evaluation, when conducted properly with valid measures and when used to improve motivation and systemic instructional practices (e.g., curriculum choices), has the potential to improve student achievement through accountability. However, summative measures generally are not useful for progress monitoring (Fuchs, 1994; Fuchs & Fuchs, 1999, 2002; D. Fuchs, L. S. Fuchs, Benowitz, & Berringer, 1987). In addition, these summative tests lack sufficient alternate forms and frequently are expensive in terms of cost and lost instructional time.

Interventions INT_Ch10_11.3d 2/2/10 19:55:36
The Charlesworth Group, Wakefield +44(0)1924 369598 - Rev 7.51n/W (Jan 20 2003)

Specified, Continuous, and Adequate Progress Monitoring

The preferred progress monitoring system within a multitiered CEIS such as RTI is *specified* explicitly, is *continuous* (i.e., the same assessment tools are used across tiers), and is *adequate*, or scientifically based. Fortunately, considerable guidance in the selection of scientifically based progress monitoring tools is available (Fuchs & Fuchs, 2004). The U.S. Department of Education Office of Special Education Programs (OSEP) funded the National Center for Student Progress Monitoring (NCSPM) for a 5-year period beginning in 2003 to support the identification and dissemination of scientifically based progress monitoring tools. Given the importance of scientifically based progress monitoring within an RTI paradigm, the NCSPM was incorporated into the OSEP's National Center on Response to Intervention (NCRTI) in 2008.

As part of their mission, the NCSPM published a set of standards for frequent progress monitoring tools based on the Standards for Educational and Psychological Testing developed by the joint committee appointed by the American Educational Research Association (AERA), the American Psychological Association (APA), and the National Council on Measurement Used in Education (NCMUE) and the Individuals with Disabilities Education Act (IDEA). According to these standards, scientifically based progress monitoring tools must (a) be reliable, generating accurate data through evidence of high parallel form and test reliability; (b) be valid, in that they measure the achievement construct of interest, including concurrent and predictive validity; (c) provide evidence of at least nine alternate forms of equivalent difficulty; (d) be sensitive to student improvement after short periods of time; (e) be linked to benchmarks specifying adequate yearly progress; (f) specify rates of improvement for typically developing students and NCLB subgroups; and (g) show evidence that the use of the progress monitoring tool results in changes in teacher instructional planning and improves student achievement. Schools that use progress monitoring tools that meet these standards are able to build decision-making practices that are explicit and specific, are continuous across tiers, and are scientifically based.

OVERVIEW OF CURRICULUM-BASED MEASUREMENT: A SET OF SCIENTIFICALLY BASED PROGRESS MONITORING TOOLS

There is widespread support for curriculum-based measurement (CBM) as a scientifically based progress monitoring tool. For example, the Review of Progress Monitoring Tools chart, now available at the National Center on Response to Intervention website, indicates that most of the tools meeting standards are types of CBM (Deno, 1985, 1995, 2002, 2003; Deno, Mirkin, & Chiang, 1982; Fuchs & Deno, 1991, 1994; Fuchs, Fuchs, Hamlett, Walz, & Germann, 1993; Fuchs & Fuchs, 1986a, 1992, 1999, 2008; Fuchs, Fuchs, Hosp, & Jenkins, 2001; Shinn, 1989, 1998, 2008). CBM also was reviewed favorably for use in progress monitoring as part of the Reading First assessment evaluation process (Kame'enui, 2002).

Interventions INT_Ch10_11.3d 2/2/10 19:55:36
The Charlesworth Group, Wakefield +44(0)1924 369598 - Rev 7.51n/W (Jan 20 2003)

CBM has been in use for almost three decades in public schools. Its initial use was as a tool to assess the progress of students with disabilities toward their IEP basic skills goals and to judge the effectiveness of their special education instructional programs (Germann & Tindal, 1985; Marston & Magnusson, 1985). Beginning in the early 1980s, the use of CBM quickly expanded to individual screening and entitlement decisions for students referred for special education (Marston, Mirkin, & Deno, 1984; Tindal & Germann, 1985).

Features and History of CBM

CBM is a family of assessment instruments that are designed to assess basic skills progress using tests with several common features. CBM tests are (a) standardized, (b) brief (i.e., usually less than 5 minutes), (c) easy to administer and score over time, (d) technically adequate, and (e) sensitive to improvement. Originally developed by Stanley Deno and a pool of graduate students, CBM was the product of directed research to address the need for scientifically based and pragmatic ways to write IEP goals and monitor progress. For early outcomes of this research program, see Deno, Marston, and Mirkin (1982) or Deno, Mirkin, and Chiang (1982); for a historical context, see the foreword in this book.

Originally, CBM assessment probes were developed from the specific curriculum used in a school's general education classrooms. However, this curriculum-specific approach, although defensible, was pragmatically challenging given curriculum differences within and between schools and over time. More importantly, this curriculum-specific approach generated test probes that ranged in difficulty level because of the inherent variability in the source curricula. Subsequent research found that alternate form reliability could be increased through use of standard probes, with further gains in logistics and without a loss in validity of progress decisions (Fuchs & Deno, 1992; Hintze & Shapiro, 1997; Hintze, Shapiro, & Lutz, 1994).

As a result, *standardized* CBM test materials that assessed basic skills in *general* rather than specific curriculum became the model of practice with different publishers of CBM materials, like any achievement test. For example, DIBELS (Dynamic Indicators of Basic Early Literacy; Kaminski & Good, 1996, 1998) is a publisher of CBM materials with an emphasis on early reading skills. Similarly, AIMSweb is a publisher of CBM materials in a variety of basic skills areas, and Monitoring Basic Skills Progress (MBSP) and Yearly Progress Pro (Fuchs, Fuchs, & Hamlett, 1995) are publishers of CBM reading and math test materials.

CBM Reading and Other Basic Skill Measures

In the area of reading, two CBM measures are the most common: (a) a 1-minute oral reading test (reading curriculum-based measurement; R-CBM) and (b) a 3- to 5-minute silent reading test (Maze). Additionally, there are a variety of early reading

Interventions INT_Ch10_11.3d 2/2/10 19:55:36
The Charlesworth Group, Wakefield +44(0)1924 369598 - Rev 7.51n/W (Jan 20 2003)

CBM measures, including Letter Naming Fluency (LNF), Letter Sound Fluency (LSF), Phonemic Segmentation Fluency (PSF), and Nonsense Word Fluency (NWF). They are similar to other CBM measures in terms of their test construction and administration features (e.g., short tests, reliability, and validity). However, they differ from the other CBM measures in that these early literacy measures are used in a mastery monitoring approach. See Fuchs and Deno (1991) for more detail on the differences between mastery monitoring and general outcome measurement.

In R-CBM, a student reads standard passages aloud, and the number of words read correctly (WRC) is counted. In contrast to popular interpretation, it would not be accurate to characterize R-CBM only as a measure of oral reading fluency (Samuels, 2007). The construct measured when students read aloud is *general reading ability* or word-level reading ability (Fuchs et al., 2001; Shinn, Good, Knutson, Tilly, & Collins, 1992). Extensive reviews of the technical adequacy of R-CBM have been published over the previous two decades (Good & Jefferson, 1998; Marston, 1989; Miura Wayman, Wallace, Ives Wiley, Ticha, & Espin, 2007).

Maze is another measure of general reading ability within the family of curriculum-based measurement (Fuchs & Fuchs, 1992; Fuchs, Fuchs, Hamlett, & Ferguson, 1992; Shin, Deno, & Espin, 2000; Shinn & Shinn, 2003). In maze, students read a passage silently and select from three choices one word that correctly preserves the meaning of the passage. The number of correct choices is counted. Although maze measures the same construct as R-CBM, it has the advantage of being administered in small or large group settings, so that economies of time can be obtained with older students (i.e., higher than Grade 3).

Although there is less published research on them, there are mature CBM tools in other basic skill areas with a history of use by schools for progress monitoring and universal screening since the late 1970s. Their use as scientifically based progress monitoring tools is evaluated annually by the NCRTI's Progress Monitoring Technical Review Committee. These CBM tests include spelling, written expression, mathematics, and early numeracy. (For spelling, see Deno, Mirkin, Lowry, & Kuehnle, 1980; Fuchs, Fuchs, Hamlett, & Allinder, 1991; for written expression, see Espin et al., 2000; McMaster & Espin, 2007; for mathematics computation and mathematics application and problem solving see Foegen, Jiban, & Deno, 2007; Fuchs, Fuchs, et al., 1994; Fuchs, Fuchs, Karns, Hamlett, & Katzaroff, 1999; Thurber & Shinn, 2002; for early numeracy see Clarke & Shinn, 2004.

DEVELOPMENT OF PROGRESS MONITORING SYSTEMS ACROSS TIERS: WORKING BACKWARD

To develop a high-quality data system for multitiered early-intervening services, including RTI, it is recommended that schools proceed "backward," beginning with progress monitoring in Tier 3, then in Tier 2, and after doing this, only then identify their universal screener. This backward approach contrasts with conventional practice, in which universal screening tools are selected first, and only then are progress

Interventions INT_Ch10_11.3d 2/2/10 19:55:36
The Charlesworth Group, Wakefield +44(0)1924 369598 - Rev 7.51n/W (Jan 20 2003)

monitoring tools identified for use at Tier 1, then Tier 2, then Tier 3. A comparison of two different approaches for creating a scientifically based progress monitoring and universal screening data system is shown in Figure 1. The left column illustrates the conventional sequence, in which reading achievement tools commonly employed in general education classrooms are used first as universal screeners, then as general education progress monitoring tools, with the presumption that these tools can be used for progress monitoring at other tiers.

Steps 1 and 2. Ensure Quality Progress Monitoring for Students Who Need It and Are Entitled to It

The right column in Figure 1 illustrates the preferred "backward" or reverse sequence in which schools develop their data system, beginning with progress monitoring for students who have a reading IEP goal (Fuchs & Fuchs, 2008; Shinn & Shinn, 2000). This sequence is recommended because CBM meets standards for scientifically based progress monitoring tools, including a demonstrated impact on student achievement. Published studies report effect sizes of .5 and greater when CBM is used to identify individualized goals and to monitor progress for students with disabilities (Fuchs & Fuchs, 1986b, 2004).

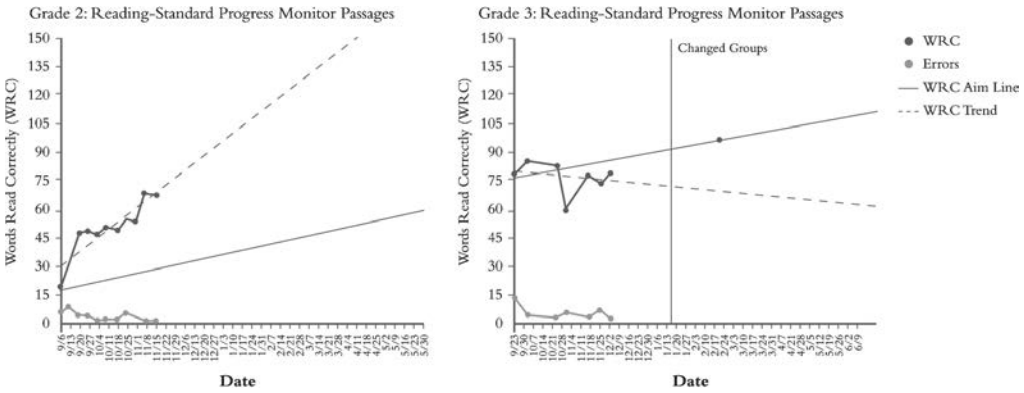
Figure 2 shows an example of two special education students whose progress on reading IEP goals was monitored using CBM. Each dot on these two graphs represents one R-CBM score on a given day. The first graph (Grade 2) demonstrates that the student's rate of improvement (trend line) exceeds the target goal (shown by the aim line),

Figure 1. A comparison of common and recommended sequences for developing a scientifically based progress monitoring and universal screening data system.

Most Common Sequence of Building RTI Data Systems Across Tiers	Recommended Sequence of Building RTI Data Systems Across Tiers
Identify Universal Screener	Step 1. Tier 3 Progress Monitoring Tool
↓	↓
Tier 1 Progress Monitoring Tool	Step 2. RTI Entitlement Progress Monitoring Tool
↓	↓
Tier 2 Progress Monitoring Tool	Step 3. Identify Universal Screener
↓	↓
Tier 3 Progress Monitoring Tool	Step 4. Tier 1 Progress Monitoring Tool
↓	↓
RTI Entitlement Progress Monitoring Tool	Step 5. Tier 2 Progress Monitoring Tool

Interventions INT_Ch10_11.3d 2/2/10 19:55:36
The Charlesworth Group, Wakefield +44(0)1924 369598 - Rev 7.51n/W (Jan 20 2003)

Figure 2. An illustration of progress monitoring toward two students' IEP goals, using CBM.



Note. From *AIMSweb.* Copyright © 2008 NCS Pearson, Inc. Reproduced with permission. All rights reserved. "*AIMSweb*" is a trademark, in the US and/or other countries, of Pearson Education, Inc. or its affiliates(s). Figure originally printed in "Best Practices in Curriculum-Based Measurement and Its Use in a Problem-Solving Model," by M. R. Shinn, 2008, in A. Thomas & J. Grimes (Eds.), *Best Practices in School Psychology V* (p. 251), Bethesda, MD: National Association of School Psychologists. Copyright 2008 by the National Association of School Psychologists.

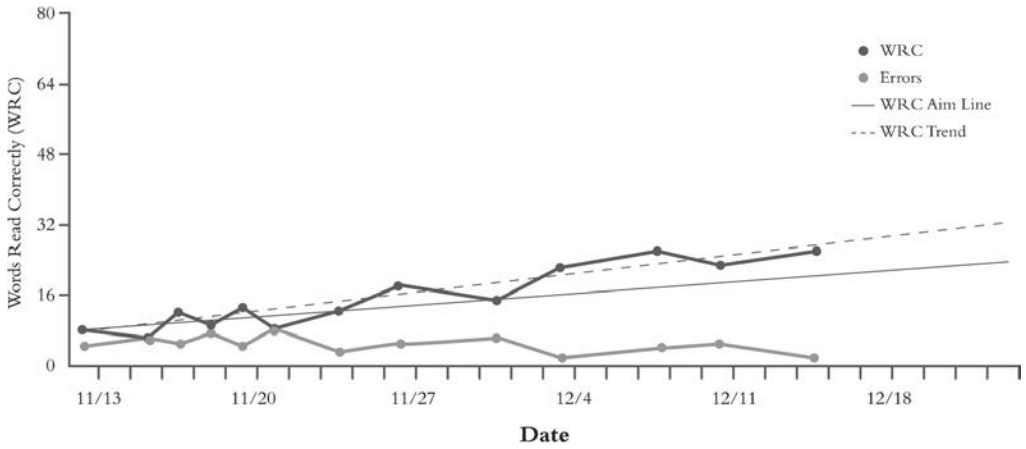
indicating that the IEP goal has been met and intervention should continue. In contrast, the graph on the right (Grade 3) illustrates the opposite pattern (i.e., the student's rate of progress is below the aim line, indicating that the intervention should be revised in line with the 1997 IDEA reauthorization's IEP requirements (Shinn & Shinn, 2000).

In the backward-design approach, CBM provides symmetry to progress monitoring practices. The same progress monitoring tool that is used to identify IEP goals and to monitor progress *after* special education eligibility (i.e., CBM; Step 1 in the recommended sequence shown in Figure 1) also guides goal setting and progress monitoring *prior to* entitlement (Step 2 in Figure 1). Both before and after the special education entitlement process, the steps are the same: (a) valid progress monitoring measures are used, (b) individualized goals are written and represented on a graph, (c) data are collected over time, (d) the rate of progress is calculated and compared with an expected rate of progress, and (e) a judgment is made about progress and response to the intervention. The only major difference between Steps 1 and 2 is the period over which progress monitoring occurs. For an IEP, the time frame is for an annual goal, whereas for an entitlement decision within RTI, the goal may be written for a much shorter period (e.g., 6–10 weeks). Figure 3 illustrates a student whose progress is monitored twice per week over a 6-week period using an R-CBM probe to determine response to a high-quality reading intervention compared with a prescribed standard of adequate progress.

Step 3. Adopt the Quality Progress Monitoring Tool as a Universal Screener

As shown in Step 3 of Figure 1, the tool (i.e., CBM) used for progress monitoring toward IEP goals and as part of the RTI entitlement decision is identified as the general

Figure 3. An illustration of progress monitoring using R-CBM as part of an RTI entitlement decision.



Note. From *AIMSweb*. Copyright © 2008 NCS Pearson, Inc. Reproduced with permission. All rights reserved. "*AIMSweb*" is a trademark, in the US and/or other countries, of Pearson Education, Inc. or its affiliates(s). Figure originally printed in "Identifying Students at Risk, Monitoring Performance, and Determining Eligibility Within RTI: Research on Educational Need and Benefit From Academic Intervention," by M. R. Shinn, 2007, *School Psychology Review, 36*, pp. 601–617. Copyright 2007 by the National Association of School Psychologists.

education universal screener. All students are screened using R-CBM; for older students, maze may be substituted. Each student's score is compared with the cut score or criterion. An illustration of how R-CBM and Maze is used with a student as part of a triage universal screening model is shown in Figure 4.

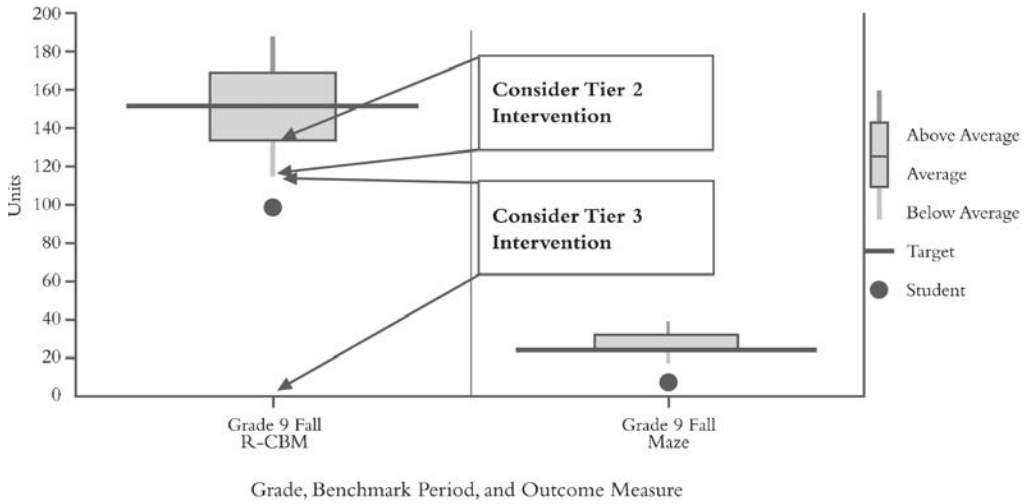
In this example, the dot represents an individual student's R-CBM score from the fall universal screening. In multitiered early-intervening services, it is recommended that a grade-level team review the screening data to match the intervention intensity with the severity of the reading problem. For example, using a normative approach, students who score below average (e.g., 10th–25th percentile) may be recommended for a Tier 2 intervention, and those who score well below average (e.g., below the 10th percentile) may be recommended for a Tier 3 intervention, as in this figure.

Step 4. Expand Universal Screening to a Benchmark Approach for Progress Monitoring

Proceeding with the recommended sequence for developing a data system (illustrated in Figure 1), the universal screening process using R-CBM then is expanded into a *benchmark* approach to provide a feasible general education progress monitoring system (Shinn, 2008). In a benchmark assessment approach, the universal screening method is repeated three to four times per year, as shown in Figure 5.

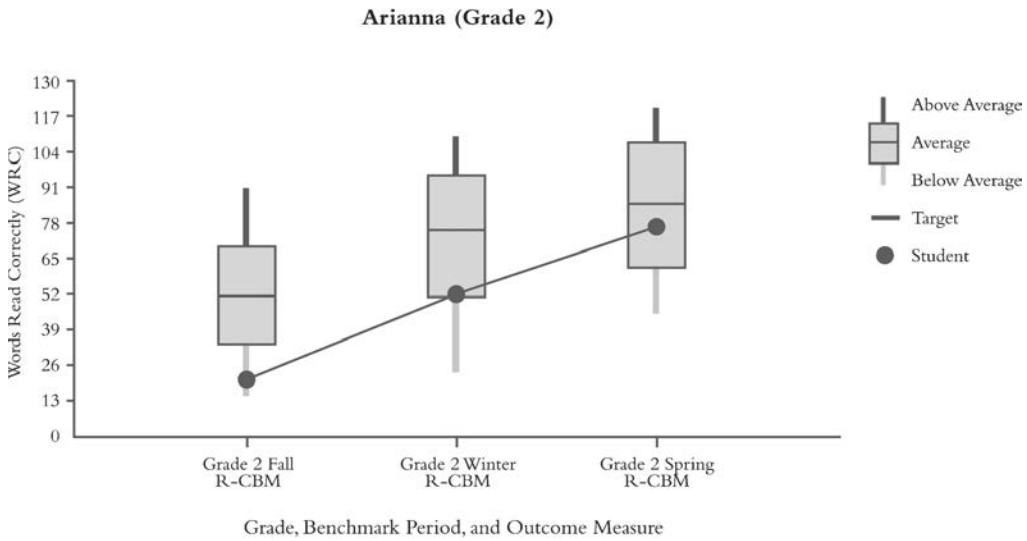
Following the school's universal screening process (shown as the fall benchmark period), the student, Arianna, was provided a Tier 2 intervention, in addition to

Figure 4. An example of R-CBM and Maze used in universal screening in a triage approach.



Note. From *AIMSweb*. Copyright © 2008 NCS Pearson, Inc. Reproduced with permission. All rights reserved. "*AIMSweb*" is a trademark, in the US and/or other countries, of Pearson Education, Inc. or its affiliates(s).

Figure 5. Progress monitoring using R-CBM for a student in Tier 1 using a benchmark approach.



Note. From *AIMSweb*. Copyright © 2008 NCS Pearson, Inc. Reproduced with permission. All rights reserved. "*AIMSweb*" is a trademark, in the US and/or other countries, of Pearson Education, Inc. or its affiliates(s). Figure originally printed in "Best Practices in Curriculum-Based Measurement and Its Use in a Problem-Solving Model," by M. R. Shinn, 2008, in A. Thomas & J. Grimes (Eds.), *Best Practices in School Psychology V* (p. 248), Bethesda, MD: National Association of School Psychologists. Copyright 2008 by the National Association of School Psychologists.

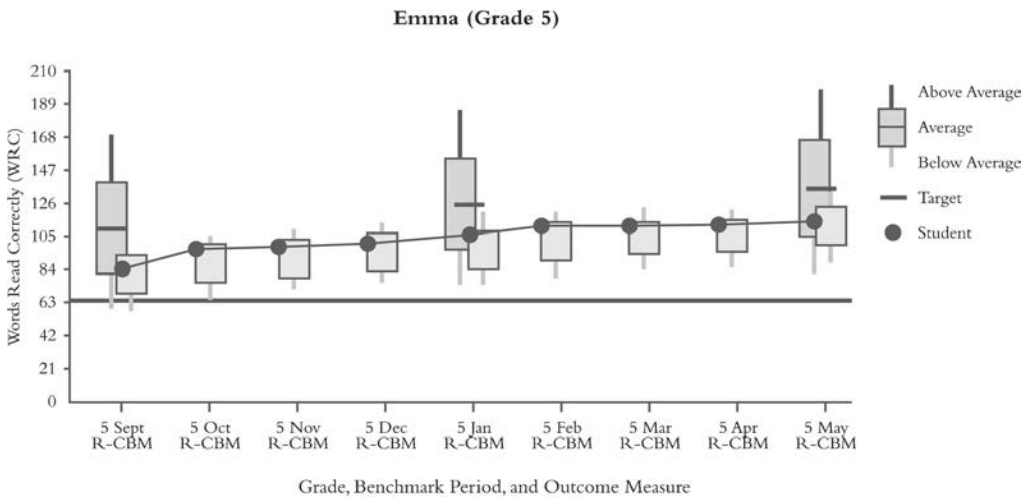
core Tier 1 instruction, because her R-CBM score was below average. By the winter benchmark period, the gap had been reduced, and she no longer required the Tier 2 intervention. Spring benchmark results illustrated that Arianna maintained an adequate rate of progress. When CBM is used in a benchmark report and the information is provided to parents, it meets the IDEA 2004 requirements for ''data-based documentation of repeated assessments of achievement at reasonable intervals, reflecting formal assessment of student progress during instruction, which was provided to the child's parents'' (34 C.F.R. 300.304–300.306).

Step 5. Ensure Continuous Progress Monitoring by Specifying Frequency for Tier 2

The final step in the recommended sequence for developing a multitiered progress monitoring system using CBM is to identify the progress monitoring strategies for students who receive a Tier 2 intervention. An example of one approach is illustrated in Figure 6.

Emma, a fifth grader, received a Tier 2 intervention in addition to her core program after the universal screening was completed using the fall benchmark assessment. In this instance, an approach called *strategic monitoring* was used to monitor her progress more frequently (i.e., once per month) than students who

Figure 6. Progress monitoring using R-CBM more frequently for a student receiving a Tier 2 intervention using a strategic monitoring approach.



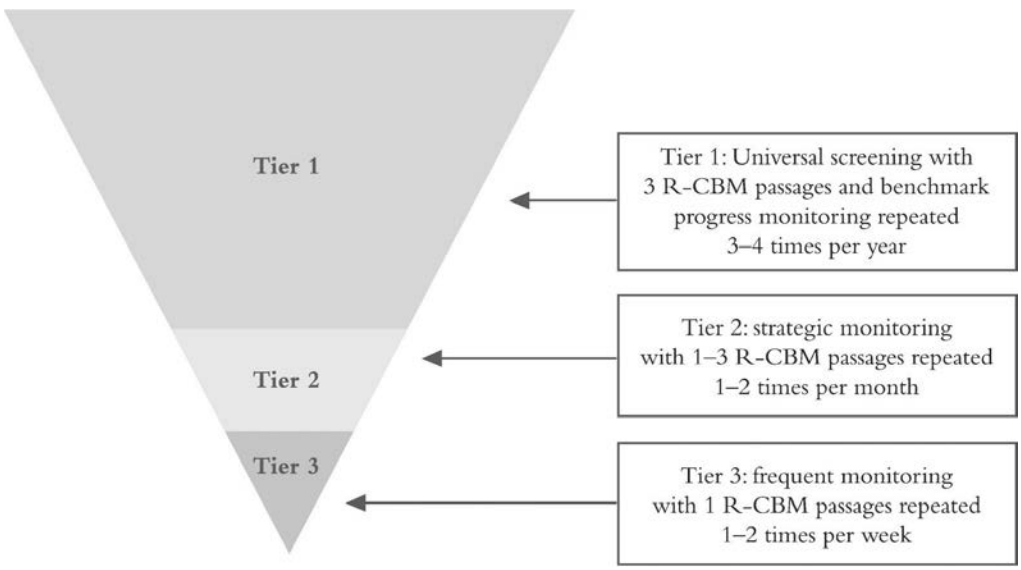
Note. From *AIMSweb*. Copyright © 2008 NCS Pearson, Inc. Reproduced with permission. All rights reserved. ''*AIMSweb*'' is a trademark, in the US and/or other countries, of Pearson Education, Inc. or its affiliates(s). Figure originally printed in ''Best Practices in Curriculum-Based Measurement and Its Use in a Problem-Solving Model,'' by M. R. Shinn, 2008, in A. Thomas & J. Grimes (Eds.), *Best Practices in School Psychology V* (p. 250), Bethesda, MD: National Association of School Psychologists. Copyright 2008 by the National Association of School Psychologists.

receive Tier 1 instruction only (three times per year). However, her progress was monitored *less frequently* than students who receive Tier 3 intervention (i.e., one to two times per week), because her reading problem was less severe than theirs. The frequency of Tier 2 progress monitoring varies from once per month to once per week, depending on the resources available.

How to Put It All Together

When a data system is designed in this way using CBM, it provides a clearly specified, continuous, and adequate (i.e., scientifically based) process for making efficient and effective decisions within a multitiered CEIS model including RTI. A graphic organizer of this data system, which illustrates when universal screening and increasingly intense progress monitoring practices are used, and for which students, is depicted in Figure 7. In contrast to many current practices, a progress monitoring database is generated within and across years, minimizing the need to collect new data should there be a concern about an individual student's progress.

Figure 7. A three-tier model for universal screening and progress monitoring for R-CBM.



Note. From "Best Practices in Curriculum-Based Measurement and Its Use in a Problem-Solving Model," by M. R. Shinn, 2008, in A. Thomas & J. Grimes (Eds.), *Best Practices in School Psychology V* (p. 247), Bethesda, MD: National Association of School Psychologists. Copyright 2008 by the National Association of School Psychologists. Adapted with permission.

SPECIFIC UNIVERSAL SCREENING AND PROGRESS MONITORING PRACTICES AND ISSUES IN USING CBM IN A MULTITIERED MODEL

In this section, the use of CBM for universal screening and progress monitoring at each tier is described, and common issues of controversy are identified.

Tier 3 Goals and Frequent Progress Monitoring

When intensive reading needs are identified through a *triage* approach to universal screening (e.g., a student reads below the 10th percentile), a student may immediately receive a Tier 3 intervention. In addition, when students do not respond to a Tier 2 intervention in a *titration* model, they may receive a Tier 3 intervention. When students have severe reading needs and do not respond to a high-quality Tier 3 intervention, they may be entitled to special education, provided eligibility criteria have been met. In each of these instances, students should be entitled to frequent progress monitoring of their response to the reading intervention using CBM with individualized goals. The process for identifying individualized goals and conducting progress monitoring using CBM is described in this section.

Practices

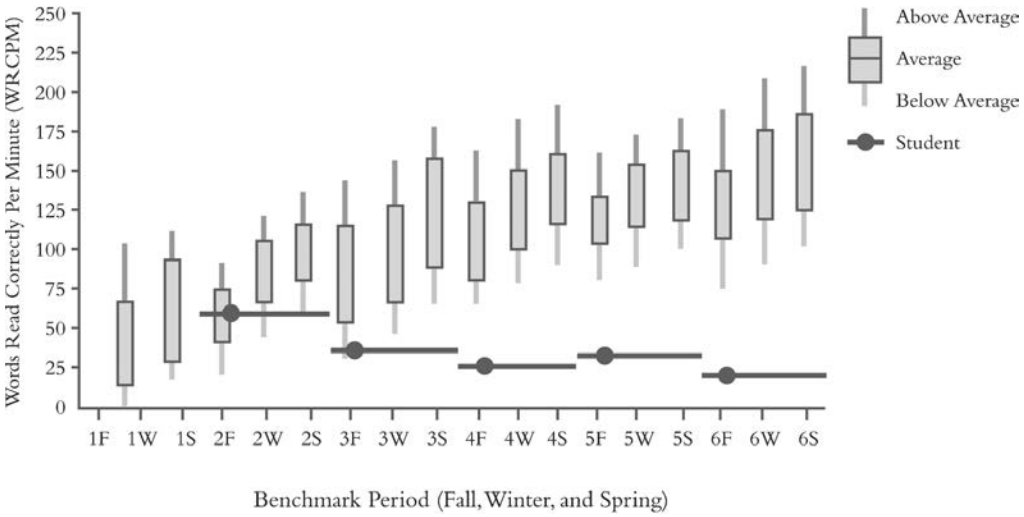
Annual goals are written based on a student's performance on graded R-CBM probes using a process called survey-level assessment (SLA). In an SLA, a student reads three passages at consecutive levels of the curriculum until he or she reads successfully, as defined by a normative score comparing performance with that of other students on those same passages. For example, as illustrated in Figure 8, Carlos, a sixth grader, read passages successively beginning at Grade 6, to Grades 5, 4, 3, and 2, the last of which he read normatively as well as other Grade 2 students. An annual goal for this student might read, ''In 1 year, when given a randomly sampled passage from the Grade 4 reading passages, Carlos will read 125 words correctly with 4 or fewer errors.''

When writing an individualized goal, the team must consider whether it is plausible for the student to eliminate the gap (i.e., be successful in Grade 6 material) in 1 year. If it is determined that this is not plausible, a lower goal would be written, but one that would be expected to reduce the gap (e.g., to read Grade 4 material successfully). For example, whereas Carlos was expected to be reading at 125 WRC in 1 year, another Grade 6 student with a more severe reading performance discrepancy may be expected to read 100 WRC from Grade 3 passages in 1 year. For more detail regarding the use of CBM to write IEP goals, the reader is referred to Fuchs and Shinn (1989); Deno, Mirkin, and Wesson (1984); Shinn (2003a); and Shinn and Shinn (2000).

CBM progress monitoring uses a general outcome measurement approach, which samples student performance on a standardized task of consistent difficulty, over time,

Interventions INT_Ch10_11.3d 2/2/10 19:56:05
The Charlesworth Group, Wakefield +44(0)1924 369598 - Rev 7.51n/W (Jan 20 2003)

Figure 8. A survey-level assessment for a Grade 6 student with severe reading needs.



Note. From *AIMSweb.* Copyright © 2008 NCS Pearson, Inc. Reproduced with permission. All rights reserved. "*AIMSweb*" is a trademark, in the US and/or other countries, of Pearson Education, Inc. or its affiliates(s). Figure originally printed in "Best Practices in Curriculum-Based Measurement and Its Use in a Problem-Solving Model," by M. R. Shinn, 2008, in A. Thomas & J. Grimes (Eds.), *Best Practices in School Psychology V* (p. 253), Bethesda, MD: National Association of School Psychologists. Copyright 2008 by the National Association of School Psychologists.

to determine if adequate growth is occurring (Fuchs & Deno, 1991; Fuchs & Shinn, 1989; Shinn, 2003a; Shinn & Shinn, 2000). For progress monitoring purposes, a student reads a single randomly sampled R-CBM passage from the goal material once or twice per week. For Carlos, this would be one or two passages per week from Grade 4 passages. As illustrated in Figure 9, a line drawn from a student's first R-CBM score (i.e., the current level of performance) to the criterion score at the goal date (i.e., the aim line) reflects adequate progress.

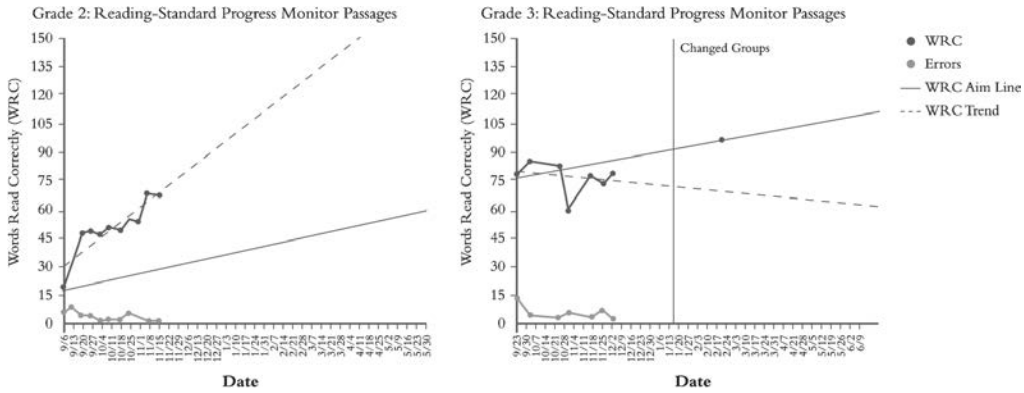
For students who are receiving Tier 3 interventions but who are not eligible for special education, it is highly desirable to monitor their progress in their grade-level material. That is, a Grade 3 student receiving a Tier 3 intervention would have a goal to be successful in Grade 3 materials. However, this decision would be made following the completion of an SLA. Figure 10 presents SLA results from a student identified as a candidate for Tier 3 intervention based on R-CBM scores below the 10th percentile. Results show that although the student is below the 10th percentile compared with other students at the same grade (i.e., that the scores are below the line representing below average), the student reads Grade 2 material successfully. Thus, the goal is Grade 3 successful reading, and progress is monitored one to two times per week using Grade 3 passages.

Controversies

Two major controversies are associated with the use of CBM for goal setting and progress monitoring at Tier 3. The first controversy is a distinct gap between the knowledge base

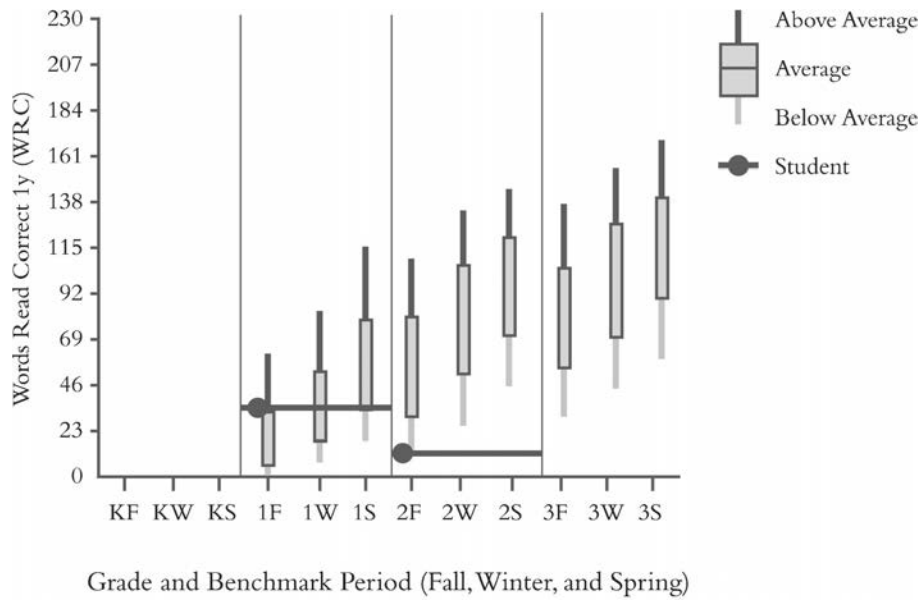
Interventions INT_Ch10_11.3d 2/2/10 19:56:05
The Charlesworth Group, Wakefield +44(0)1924 369598 - Rev 7.51n/W (Jan 20 2003)

Figure 9. An illustration of progress monitoring toward two students' IEP goals, using CBM.



Note. From *AIMSweb*. Copyright © 2008 NCS Pearson, Inc. Reproduced with permission. All rights reserved. "*AIMSweb*" is a trademark, in the US and/or other countries, of Pearson Education, Inc. or its affiliates(s). Figure originally printed in "Best Practices in Curriculum-Based Measurement and Its Use in a Problem-Solving Model," by M. R. Shinn, 2008, in A. Thomas & J. Grimes (Eds.), *Best Practices in School Psychology V* (p. 251), Bethesda, MD: National Association of School Psychologists. Copyright 2008 by the National Association of School Psychologists.

Figure 10. A survey-level assessment for a student identified for potential Tier 3 services using R-CBM.



Note. From *AIMSweb*. Copyright © 2008 NCS Pearson, Inc. Reproduced with permission. All rights reserved. "*AIMSweb*" is a trademark, in the US and/or other countries, of Pearson Education, Inc. or its affiliates(s). Figure originally printed in "Best Practices in Curriculum-Based Measurement and Its Use in a Problem-Solving Model," by M. R. Shinn, 2008, in A. Thomas & J. Grimes (Eds.), *Best Practices in School Psychology V* (p. 253), Bethesda, MD: National Association of School Psychologists. Copyright 2008 by the National Association of School Psychologists. Reprinted with permission.

and standard practice. That is, despite long-standing concerns over the quality of IEP goals and progress monitoring practices (Fuchs, Deno, & Mirkin, 1982; Smith, 1990), these have remained largely unchanged since 1975 (Bateman & Linden, 1998; Giangreco, Dennis, Edelman, & Chigee, 1994; Shinn & Shinn, 2000). This gap between scientifically based IEP progress monitoring practices and what typically occurs in schools remains wide, despite professionals' familiarity with CBM (Shapiro, Angello, & Eckert, 2004) and the wide availability of excellent resources (e.g., NCRTI).

A second controversy concerns the goal-setting process. For example, although at least four possible methods for setting a criterion for success have been identified, not a single comparative study of these approaches has been conducted. More guidance about goal setting is needed (Fuchs, Fuchs, & Deno, 1985).

Progress Monitoring and RTI as an Entitlement Process

In assessing RTI as an SLD entitlement process, a *dual-discrepancy* approach has been recommended (Fuchs, Fuchs, & Speece, 2002; Pericola Case et al., 2003). A student who exhibits a dual discrepancy has (a) severe low achievement relative to peers (i.e., discrepancy in educational need), which can be assessed economically and accurately using R-CBM data from universal screening; and (b) a low rate of improvement, which can be determined from progress monitoring data showing a lack of response to appropriate instruction (discrepancy of educational benefit; Shinn, 2007).

Practices

Two progress monitoring approaches can be used in an RTI entitlement process: (a) using extant CBM data from frequent progress monitoring, and (b) developing a new progress monitoring plan. For a student who has been receiving an appropriate (i.e., high quality, high intensity) Tier 3 intervention along with CBM progress monitoring, these data provide the basis for the judgment of appropriate response to intervention. In the second instance, if a student has not received a Tier 3 intervention and is considered for special education entitlement, then a short-term goal (e.g., 6 weeks) would need to be written and progress would be monitored during the intervention period.

An example of progress monitoring as part of RTI entitlement for a Grade 2 student is shown in Figure 3. The goal is written so that if the goal were attained, the gap in reading Grade 2 material, compared with peers, would be reduced. It is preferable that a sample of at least 10 data points be obtained. If the goal was written for a 6-week time frame and student progress was monitored twice per week, a judgment of adequate progress would be based on 12 data points. If the goal were written for an 8-week time frame, a judgment of adequate progress would be based on 16 data points.

Controversies

If schools use CBM to write quality IEP goals and monitor progress for students with disabilities, extending these practices as described in this section for RTI entitlement is

Interventions INT_Ch10_11.3d 2/2/10 19:56:18
The Charlesworth Group, Wakefield +44(0)1924 369598 - Rev 7.51n/W (Jan 20 2003)

reasonably simple. However, if scientifically based progress monitoring using CBM is not standard practice within a district's special education program, schools will need to add this capacity in addition to the numerous challenges associated with developing intensive multitiered coordinated early-intervening services.

In addition, there is not a consensus regarding appropriate goal-setting practices within the RTI entitlement process. While some would assert that adequate progress is defined as a rate of improvement equivalent to typically developing peers, others argue that adequate progress must reduce the achievement gap. Further research to address this question would assist practitioners in using CBM for RTI entitlement.

CBM and Universal Screening

Since the early 1980s, CBM has been used to screen individual students for potential reading difficulties (Marston, Deno, & Tindal, 1984; Marston, Mirkin, & Deno, 1984). However, this screening typically began with special education referral one student at a time. Getting *all* students off to a good reading start begins by implementing a universal screening process in which *all* students are tested to identify potentially at-risk students. CBM is a time-efficient tool for this process. More important, there is an accumulating body of knowledge supporting the validity of using CBM in universal screening (Stage & Jacobsen, 2001). In universal screening, two measures of predictive efficiency are most important: (a) *sensitivity*, or true positives (i.e., those students predicted to not pass the criterion test who in fact do not pass), and (b) *specificity*, or true negatives (i.e., those students predicted to pass the criterion test who do pass). Two recent studies that compared the diagnostic accuracy and predictive validity of R-CBM with high-stakes state reading tests reported ranges of predictive efficiency of .65 to .76 for sensitivity and .78 to .82 for specificity, with predictions from the winter Grade 1 R-CBM and the end of Grade 3 state reading test (Hintze & Silberglitt, 2005; Silberglitt & Hintze, 2005).

Practices

When R-CBM is used in universal screening, students typically read three graded passages, and the median WRC is counted. This process is completed in approximately 5 minutes per student when the examiners are well trained. Many CBM publishers also offer options for collecting, scoring, and reporting options using hand-held electronic devices that can increase efficiency further. With older students (e.g., Grade 5 and above), some economies of time can be achieved by substituting maze—a 3- to 5-minute silent reading test that can be administered to students in groups through pencil-and-paper testing or through use of a computer—instead of R-CBM (Stage & Jacobsen, 2001).

As discussed earlier in the chapter, universal screening practices most often are based on the titration approach, in which at-risk students are identified for consideration for Tier 2 interventions. Progress through the remaining tiers and

potential special education entitlement is based on CBM progress monitoring data. In the titration approach, the first student on the left in Figure 11 would receive the core reading program through Tier 1 instruction, while both students to the right would be considered for Tier 2 intervention.

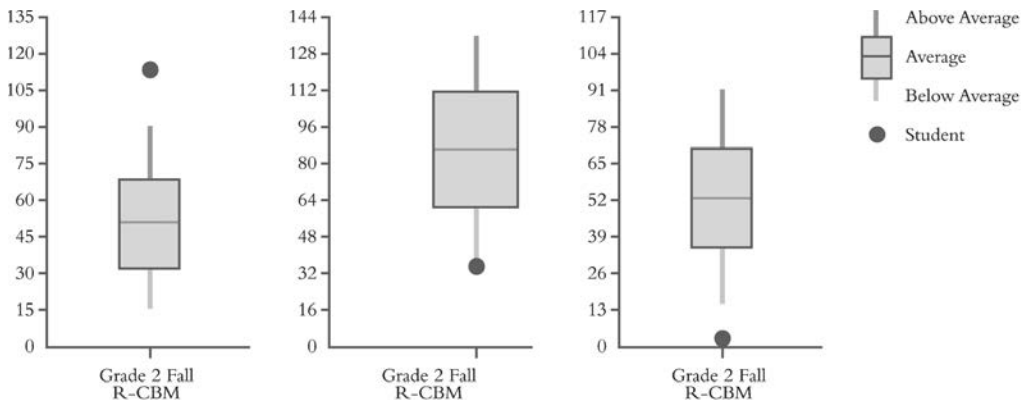
Increasingly, however, schools are using the triage approach, in which interventions other than Tier 2 are considered, based on the severity of the reading problem identified through universal screening. Students performing below average (e.g., 10th to the 25th percentile) may be recommended for Tier 2 intervention. Students well below average (e.g., below the 10th percentile) may be recommended for intensive Tier 3 interventions. In Figure 11, the first student on the left would receive the core reading program through Tier 1 instruction. However, within a triage approach, the middle student likely would be recommended to receive a Tier 2 intervention, while the student to the right may be considered for Tier 3 intervention.

Controversies

To date, there are no studies comparing the titration and triage approaches to universal screening with respect to effects on achievement. It is plausible that the sooner a student with a severe reading problem receives the most intensive intervention, the greater the likelihood of reducing the reading achievement gap. From a social validity perspective, it is plausible that the triage approach may be judged more favorably by teachers and parents because there is no need for a student to first fail in less-intensive interventions.

Considerable controversy exists regarding how the criteria for defining students as potentially at risk are specified. There are no empirically validated standards for using

Figure 11. Results of universal screening using R-CBM in a benchmark approach for three students.



Note. From *AIMSweb*. Copyright © 2008 NCS Pearson, Inc. Reproduced with permission. All rights reserved. "*AIMSweb*" is a trademark, in the US and/or other countries, of Pearson Education, Inc. or its affiliates(s). Figure originally printed in "Best Practices in Curriculum-Based Measurement and Its Use in a Problem-Solving Model," by M. R. Shinn, 2008, in A. Thomas & J. Grimes (Eds.), *Best Practices in School Psychology V* (p. 249), Bethesda, MD: National Association of School Psychologists. Copyright 2008 by the National Association of School Psychologists.

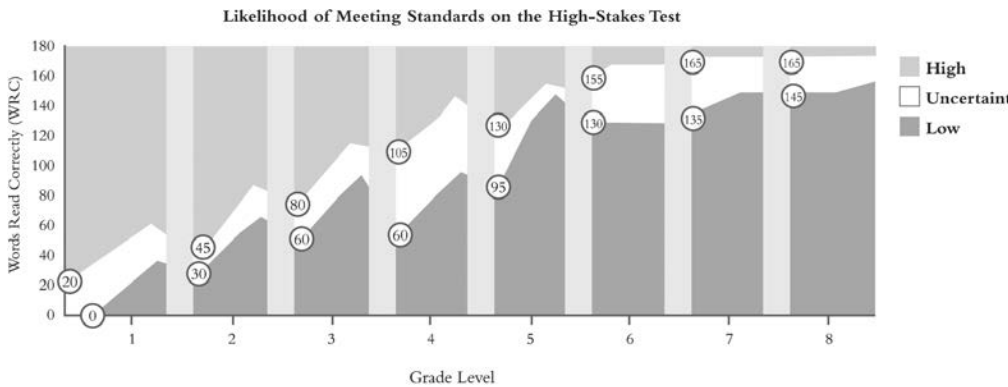
Interventions INT_Ch10_11.3d 2/2/10 19:56:19
The Charlesworth Group, Wakefield +44(0)1924 369598 - Rev 7.51n/W (Jan 20 2003)

R-CBM or any other test for this purpose. Instead, some researchers propose that such criteria are derived from the value judgments of the school personnel, based in part on the availability of intervention resources (Deno, 1989; Shinn, Good, & Parker, 1999). For example, in a district with significant resources for outside-the-classroom support, 25% of students may be served in Tier 2 interventions. The normative cut score for this setting would be those scores below the 35th percentile for Tier 2 interventions and below the 10th percentile for Tier 3. In districts with fewer outside-the-classroom support resources, it may be possible to provide Tier 2 interventions to only 15% of students. In that case, the criterion for potential Tier 2 intervention would be scores below the 25th percentile.

In part, to offset the challenges introduced by setting cut scores based on normative data (e.g., intervening with X percent of students), some schools have chosen to link their cut scores to *standards*. These standards are indeed more objective than normative cut scores based on social values and available resources, because they are derived by examining correlations of R-CBM with state high-stakes reading tests and predictive validity. Through the use of statistical methods (e.g., receiver operator characteristic [ROC] curves, logistic regression), critical values associated with high and low probabilities of passing the high-stakes test can be computed. An excellent example is shown in Figure 12, which displays critical R-CBM scores for Grades 1 to 8 fall, winter, and spring benchmarks as they relate to state standards. As illustrated in this figure, a score of 80 WRC or greater at the beginning of Grade 3 is associated with a high probability of exceeding state standards. A score of 60 WRC or lower is associated with a low likelihood of passing the state test.

When this *standards-based* approach to universal screening is used, those students who are less likely to meet or exceed state standards may be considered for a Tier 2 ''strategic'' intervention (Simmons et al., 2002). The standards-based universal

Figure 12. Standards-based cut scores across Grades 1 to 8 based on R-CBM scores and the Illinois Standards Achievement Test (ISAT).



Note. This figure was developed by Ben Ditkowsky, PhD from http://measuredeffects.com. Reprinted with permission.

Interventions INT_Ch10_11.3d 2/2/10 19:56:22
The Charlesworth Group, Wakefield +44(0)1924 369598 - Rev 7.51n/W (Jan 20 2003)

screening approach is discussed in the section on controversy rather than practices because of the unintended consequences of this approach. For example, in a low-performing school, more than 60% of the students were identified as needing Tier 2 interventions because they performed below the R-CBM standards-based cut score. Of course, the school did not have sufficient resources to deliver this magnitude of Tier 2 intervention *in addition* to their core Tier 1 program.

Although intuitively appealing, a standards-based approach is best used for *program evaluation* purposes instead of universal screening. That is, a school that has more than 60% of students being unlikely to pass the state test suggests the need to *improve the core* (i.e., Tier 1) program rather than target individual students for intervention.

IMPLEMENTATION OF PROGRESS MONITORING IN THREE TIERS USING A BENCHMARK AND STRATEGIC MONITORING APPROACH

Monitoring progress toward IEP goals to evaluate the effects of special education intervention or to evaluate RTI as part of special education entitlement requires *frequent* (e.g., at least once per week) progress monitoring. However, progress monitoring is an important component of school success for *all* students. If all students benefit from progress monitoring, the question becomes one of how to do it in a manner that is *feasible,* one of a number of important criteria in a school-based progress monitoring model (Fuchs & Fuchs, 1999). Schools have been successful in making progress monitoring feasible by using benchmark assessment for all students and by using strategic monitoring for students at risk.

Tier 1 Benchmark Assessment

Benchmark assessment, benchmark testing, or "benchmarking" are synonymous terms used to describe the process of using CBM for universal screening *and* progress monitoring. When R-CBM is used in benchmarking, the initial universal screening process is repeated two to three more times during the academic year (e.g., winter and spring). Following this practice, benchmarking consists of a fall universal screening and two subsequent progress monitoring assessments, requiring approximately 15 minutes per student per year using R-CBM, and less than 20 minutes per year for an entire grade using maze materials.

Practices

Two common approaches are used in the benchmark reading assessments. In the first approach, the same test materials (i.e., passages or maze) are used each time. Although this approach ensures that the test probes are of equal difficulty, there are potential costs involved. The first relates to possible loss of test security; since the specific passages are known in advance, it is possible that students will practice them, invalidating their

Interventions INT_Ch10_11.3d 2/2/10 19:56:25
The Charlesworth Group, Wakefield +44(0)1924 369598 - Rev 7.51n/W (Jan 20 2003)

usefulness as a progress monitoring tool. Additionally, given that the same passages are used repeatedly, there is the small likelihood of a practice effect, although no study has validated this effect during benchmarking. Finally, there is the potential problem of examiner fatigue. When examiners are routinely and frequently engaging in the same testing practices, error can be increased by this repetition over time.

The second approach for benchmarking uses alternate forms, in which different R-CBM or maze probes are used each time. This approach eliminates the effects of loss of test security, practice, and potential examiner fatigue. However, it introduces the potential disadvantage of each set of probes differing in difficulty. As a consequence, when this benchmark approach is used, it is important that the pool of potential assessment materials be studied with respect to their alternate form reliability and evidence of equal difficulty be provided.

Following the administration of R-CBM or maze probes during a benchmark period, each student's scores are graphed on a normative chart and a decision is made regarding the student's rate of improvement and current performance level. An example of benchmark assessment for a Grade 2 student over three testing periods is shown in Figure 5. The fall benchmark score for Arianna was used for purposes of universal screening. Because these data indicated her performance was below average, the grade-level team recommended a Tier 2 intervention. By the winter benchmark, Arianna's rate of improvement was above that of typical students and the achievement gap was reduced. By the spring benchmark, she read as well as other students.

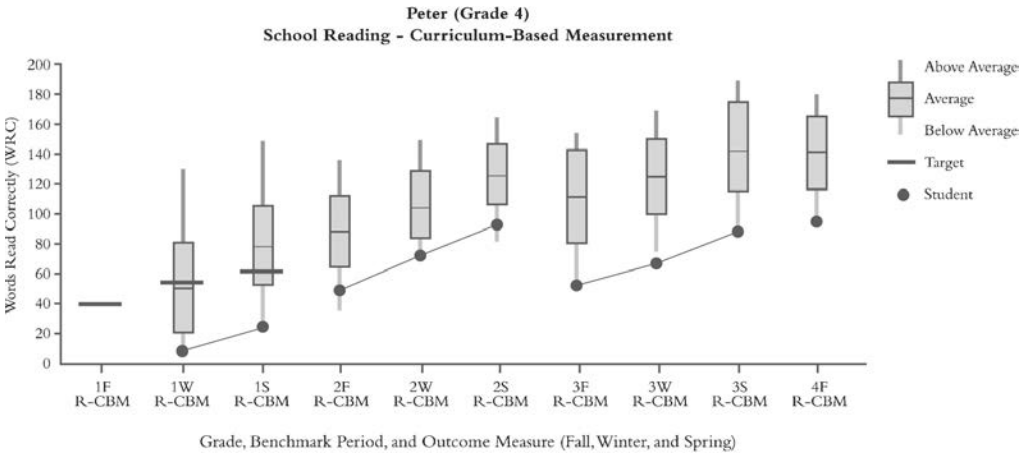
By compiling benchmark scores across years, a graphic display of reading progress from one grade to another can be created. Such a display may be particularly useful for tracking progress and reporting to parents of students receiving special education or other remedial services.

The R-CBM benchmark scores for a student from mid-Grade 1 to the beginning of Grade 4 are shown in Figure 13. Although Peter demonstrates very low reading scores, his reading improved each year and at about the same rate as typically developing peers who received the same instruction. However, if the goal was to reduce the reading skill gap, the reading program was not accomplishing its intended outcome. In Grade 4, the school moved to building a multitiered early-intervening service model.

Controversies

The principal controversies regarding benchmarking are related to ownership and logistics. With respect to the former, in a number of districts, benchmarking evolves from developing local norms. In local norming, a *subset* of students is tested to assist in screening individual students referred for special education (Stewart & Silberglitt, 2008). This subset of students often is tested by a group of special educators, and general education teachers typically are not involved. When the local norming process expands to benchmarking all students, and general education teachers are expected to participate in testing, some perceive this as more work. The second issue concerns making the benchmarking process as efficient as possible. Even an individual student 5-

Figure 13. A fourth-grade student's benchmark scores from mid-first grade to early fourth grade.



Note. From *AIMSweb*. Copyright © 2008 NCS Pearson, Inc. Reproduced with permission. All rights reserved. "*AIMSweb*" is a trademark, in the US and/or other countries, of Pearson Education, Inc. or its affiliates(s).

minute benchmark assessment compounds into hours when multiplied by the number of students tested. Every effort should be made to ensure that data are collected in an efficient manner, and specific training materials have been developed to facilitate this goal (Shinn, 2003b).

Tier 2 Progress Monitoring Using a Strategic Approach

In a multitiered model, students who receive Tier 2 interventions are provided more intensive interventions because they are at greater risk for reading failure. Because of this risk, schools also monitor their progress more frequently than that of the students who receive Tier 1 intervention alone, but typically not with the regularity of progress monitoring for students who receive Tier 3 interventions. The type of progress monitoring practices used with students at risk but without severe educational need is called *strategic monitoring*, to distinguish it from benchmarking or frequent progress monitoring.

Practices

The strategic monitoring process is similar to that of benchmarking. Students' progress is monitored using grade-level reading probes (i.e., a Grade 3 student receiving a Tier 2 intervention would be assessed using Grade 3 passages), and the individual student's goal is to achieve at the same rate as typically developing students. Strategic monitoring differs from benchmarking in the frequency of progress monitoring.

Most commonly, strategic monitoring involves monthly testing, using the three benchmark scores and repeating the benchmark assessment in the months when there is

Interventions INT_Ch10_11.3d 2/2/10 19:56:26
The Charlesworth Group, Wakefield +44(0)1924 369598 - Rev 7.51n/W (Jan 20 2003)

no benchmarking. To illustrate, in a school where benchmarking occurs in the fall for purposes of universal screening and is repeated in January and late April for progress monitoring, students receiving Tier 2 intervention would read three passages from grade-level material for strategic monitoring during months without benchmarking (e.g., October, November, December, etc.). This strategic monitoring approach requires a commitment of approximately 5 minutes per month per student. An example of a student whose progress is monitored strategically is shown in Figure 6.

In this figure, Emma's score is represented by the dot and is compared with the box and whisker plot of all the students in the school or district and with the disaggregated subset of other students who receive Tier 2 interventions. Each month that strategic monitoring takes place, the scores are plotted against other students who receive Tier 2, until the next benchmark, when the scores are compared with all students, as well as with the disaggregated subgroup of students in Tier 2. Figure 6 illustrates that Emma is making the same rate of progress as other students overall, as well as other students receiving the Tier 2 interventions.

A second, but less common approach to strategic monitoring is to use the type of progress monitoring practices in Tier 3 or for IEP progress monitoring. When this approach is used, the process is different in two regards. First, the student is tested in grade-level materials with standard goals rather than through an individualized goal-setting process. Second, progress is monitored more frequently, usually every other week or once per week.

Controversies

The primary controversies in Tier 2 progress monitoring are twofold. First, as noted in chapters 6 and 18 in this book, the real challenge in developing multitiered early-intervening services is ensuring coordinated and powerful Tier 2 interventions. Establishing a common, scientifically based progress monitoring system is part of that challenge. The second controversy concerns identifying which progress monitoring practices are most effective in increasing student achievement. To date, there have been no comparative studies examining which CBM progress monitoring practices (e.g., frequency, goal-setting practices) produce the best outcomes with at-risk students.

UNIVERSAL SCREENING AND PROGRESS MONITORING IN THREE TIERS IN EARLY LITERACY

Although some students are reading when they enter first grade, typically achieving early first graders earn very low scores on R-CBM first-grade passages. With the exception of some high-achieving communities, the distribution of R-CBM scores is positively skewed, with many students earning scores of less than 5 WRC (Rodden-Nord & Shinn, 1991). About the middle of first grade, R-CBM begins to be more useful for identifying at-risk students and for progress monitoring.

Because a major goal of reading success is early detection of students at risk for reading difficulties, waiting until most students can read R-CBM Grade 1 passages is not acceptable. This is a well-known shortcoming of R-CBM Grade 1 passages, and educators have long sought to develop other assessment tools for identifying at-risk students and for progress monitoring in kindergarten and early Grade 1.

A synthesis of the scientific research on reading by the National Reading Panel (2000) provided some critical reading skills that could (and should) be assessed in kindergarten and early Grade 1, including phonemic awareness and elements of phonics, including letter names and sounds, and the ability to read nonsense words or lists of phonetically regular words. Unfortunately, no *single* measure that can be used continuously across these early literacy skills has been validated as an early literacy progress monitoring tool. As a result, different early reading skills must be assessed at different periods.

DIBELS (Dynamic Indicators of Basic Early Literacy Skills; Kaminski & Good, 1996, 1998) was developed as a downward extension of CBM to early literacy to help solve the problem of waiting for R-CBM to be sensitive to between- and within-student reading differences. DIBELS initially consisted of these early literacy tests: Initial Sound Fluency (ISF), Phonemic Segmentation Fluency (PSF), and Nonsense Word Fluency (NWF). It was later expanded to include Letter Naming Fluency (LNF), an R-CBM task, and other measures of vocabulary and comprehension.

Although DIBELS has been referred to as a progress monitoring tool for general outcome measurement, in which a single skill can be used to monitor progress over an extended period of time, most DIBELS measures fit more closely with a mastery monitoring approach. In mastery monitoring, specific subskills are assessed at specific times in line with a logical or empirical sequence (Fuchs & Deno, 1991). With DIBELS, the measures represent a developmental hierarchy, with ISF viewed as a precursor to PSF, PSF as a precursor to NWF, and NWF as a precursor to R-CBM.

When using DIBELS, multiple measures are administered to all students at each benchmark period in kindergarten through Grade 2, both for universal screening and progress monitoring, as described in this chapter for R-CBM (see Kaminski, Cummings, Powell-Smith, & Good, 2008, for more detail on specific DIBELS practices). DIBELS differs from R-CBM in two ways. First, in universal screening, DIBELS almost always uses a standards-based approach. Reports target students whose scores fall below a criterion linked to a state test for additional intervention. Second, because implicitly it is a mastery monitoring approach, different early literacy skills must be targeted for progress monitoring at Tiers 2 and 3 at different times. For example, in early kindergarten a student receiving a Tier 2 reading intervention may first be progress monitored using PSF and then be progress monitored with NWF when the PSF criterion is attained.

In kindergarten, up to four DIBELS measures may be given to all students. In the fall of Grade 2, according to DIBELS benchmark specifications (Kaminski et al., 2008), students still are tested on NWF. As a consequence of administering multiple subskill tests to all students over time, DIBELS has been criticized for overtesting students

Interventions INT_Ch10_11.3d 2/2/10 19:56:30
The Charlesworth Group, Wakefield +44(0)1924 369598 - Rev 7.51n/W (Jan 20 2003)

(Pearson, 2006) and for being logistically challenging (Fuchs, Fuchs, & Compton, 2004).

To improve feasibility, it is suggested that three alternative approaches to early literacy universal screening and progress monitoring be considered. First, LNF might be used *only* as a fall kindergarten screener. This type of measure has a long-standing empirically demonstrated predictive relation to reading failure (Bond & Dykstra, 1967) that has been replicated in recent years (Elliott, Lee, & Tollefson, 2001; Hintze, Ryan, & Stoner, 2003). Its efficiency and accuracy make it an excellent universal screening tool. However, there is no evidence that LNF is useful as a progress monitoring tool.

Second, it is suggested that these literacy tools be used in a *multiple-gating* approach. In a multiple-gating approach, all students are tested on the *highest skill* in a hierarchical skill set. For example, if a Grade 1 student performs satisfactorily on R-CBM, a measure of general reading ability, then that student would not be tested on any of the DIBELS subskills (e.g., NWF, PSF). Similarly, a kindergarten student who performs satisfactorily on NWF would not be tested on PSF. This multiple gating approach reduces the amount of testing for all students and enables schools to collect the most information on the students who are at greatest risk.

Third, it is suggested that schools use Letter Sound Fluency (LSF) as the primary measure for universal screening *and* progress monitoring after the fall kindergarten benchmarking. LSF is another measure of the alphabetic principle, it is very easy to administer, and it works well as a universal screener. In a predictive validity study with over 2,000 students, kindergarten LSF was shown to be superior to PSF as a predictor of spring Grade 1 reading (.66 vs. .35; Silberglitt, 2007). In this same longitudinal study, LSF also demonstrated considerable sensitivity to improvement, with changes in means from 10.1 to 21.5 to 40.3 letter sounds at each benchmark. As a progress monitoring tool, LSF is highly correlated with PSF and NWF and can simplify mastery monitoring by providing a single progress monitoring test. Additionally, letter sounds are a more suitable and authentic instructional target than either PSF or NWF.

CONCLUSION

This chapter illustrated how curriculum-based measurement—a set of simple, inexpensive, time-efficient, and scientifically based basic skills tests—can be used across multiple tiers for progress monitoring and universal screening. Although CBM for readings was used as the primary exemplar, CBM applications for older students and younger students were included (e.g., maze, and modifications or additions to DIBELS).

With more than 250 published journal articles and book chapters on R-CBM alone, curriculum-based measurement is highly endorsed as an evidence-based practice for universal screening and progress monitoring. Certainly, CBM is a cornerstone of RTI, especially when RTI is conceived as multitiered early-intervening services.

Despite more than 30 years of research and applied practice, CBM remains underutilized within educational settings, which too frequently retain traditional, albeit questionable measures for screening and progress monitoring. It is hoped that this chapter, and other relevant contributions within this volume, will provide practitioners with the necessary information, skills, and resources for translating what works to widespread school practice.

REFERENCES

Adams, M. J. (1990). *Beginning to read: Thinking and learning about print*. Cambridge, MA: MIT Press.

Algozzine, B., Christenson, S., & Ysseldyke, J. (1982). Probabilities associated with the referral to placement process. *Teacher Education and Special Education, 5,* 19–23.

Baker, E. L., O'Neil, H. F., & Linn, R. L. (1993). Policy and validity prospects for performance-based assessment. *American Psychologist, 48,* 1210–1218.

Bateman, B. D., & Linden, M. A. (1998). *Better IEPs*. Longmont, CO: Sopris West.

Bond, G. L., & Dykstra, R. (1967). The cooperative research program in first-grade reading instruction. *Reading Research Quarterly, 2,* 5–14.

Burdette, P. (2007). *Response to intervention as it relates to early intervention services*. Washington, DC: Project Forum at National Association of State Directors of Special Education.

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. (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. *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., Marston, D., & Mirkin, P. (1982). Valid measurement procedures for continuous evaluation of written expression. *Exceptional Children, 48,* 68–71.

Deno, S. L., Mirkin, P., & Chiang, B. (1982). Identifying valid measures of reading. *Exceptional Children, 49,* 36–45.

Deno, S. L., Mirkin, P. K., Lowry, L., & Kuehnle, K. (1980). *Relationships among simple measures of spelling and performance on standardized achievement tests*. Minneapolis: University of Minnesota, Institute for Research on Learning Disabilities.

Deno, S. L., Mirkin, P., & Wesson, C. (1984). How to write effective data-based IEPs. *Teaching Exceptional Children, 16,* 99–104.

Elliott, J., Lee, S. W., & Tollefson, N. (2001). A reliability and validity study of the Dynamic Indicators of Basic Early Literacy Skills–Modified. *School Psychology Review, 30,* 33–49.

Espin, C., Shin, J., Deno, S. L., Skare, S., Robinson, S., & Benner, B. (2000). Identifying indicators of written expression proficiency for middle school students. *Journal of Special Education, 34,* 140–153.

Fletcher, J. M., Coulter, W. A., Reschly, D. J., & Vaughn, S. (2004). Alternative approaches to the definition of learning disabilities: Some questions and answers. *Annals of Dyslexia, 54,* 304–331.

Foegen, A., Jiban, C., & Deno, S. L. (2007). Progress monitoring measures in mathematics: A review of the literature. *Journal of Special Education, 41,* 121–139.

Fuchs, D., Fuchs, L. S., Benowitz, S., & Berringer, K. (1987). Norm-referenced tests: Are they valid for use with handicapped students? *Exceptional Children, 54,* 263–271.

Fuchs, L. S. (1994). *Connecting performance assessment to instruction*. Reston, VA: Council for Exceptional Children.

Fuchs, L. S., & Deno, S. L. (1991). Paradigmatic distinctions between instructionally relevant measurement models. *Exceptional Children, 57,* 488–500.

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. K. (1982). *Special education practice in evaluating student progress toward goals* (No. IRLD-RR-81). Minneapolis: University of Minnesota, Institute for Research on Learning Disabilities.

Fuchs, L. S., & Fuchs, D. (1984). Criterion-referenced assessment without measurement: How accurate for special education. *Remedial and Special Education, 5,* 29–32.

Fuchs, L. S., & Fuchs, D. (1986a). Curriculum-based assessment of progress towards long- and short-term goals. *Journal of Special Education, 20,* 69–82.

Fuchs, L. S., & Fuchs, D. (1986b). Effects of systematic formative evaluation on student achievement, A meta-analysis. *Exceptional Children 53,* 199–208.

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. (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. (2005). Responsiveness-to-intervention: A blueprint for practitioners, policymakers, and parents. *Teaching Exceptional Children*, Sept./Oct., 57–61.

Fuchs, L. S., & Fuchs, D. (2008). Best practices in progress monitoring reading and mathematics at the elementary level. In A. Thomas & J. Grimes (Eds.), *Best practices in school psychology V* (pp. 2147–2164). Bethesda, MD: National Association of School Psychologists.

Fuchs, L. S., Fuchs, D., & Compton, D. L. (2004). Monitoring early reading development in first grade: Word identification fluency versus nonsense word fluency. *Exceptional Children, 71,* 7–21.

Fuchs, L. S., Fuchs, D., & Deno, S. L. (1985). The importance of goal ambitiousness and goal mastery to student achievement. *Exceptional Children, 52,* 63–71.

Fuchs, L. S., Fuchs, D., & Hamlett, C. (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., Thompson, A., Roberts, P. H., & Kupek, P. (1994). Technical features of a mathematics concepts and applications curriculum-based measurement system. *Diagnostique, 19,* 23–49.

Fuchs, L. S., Fuchs, D., Hamlett, C. L., Walz, L., & Germann, G. (1993). Formulative evaluation of academic progress: How much growth can we expect? *School Psychology Review, 22*(1), 27–48.

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., Karns, K., Hamlett, C. L., & Katzaroff, M. (1999). Mathematics performance assessment in the classroom: effects on teacher planning and student problem solving. *American Educational Research Journal, 36,* 609–646.

Fuchs, L. S., Fuchs, D., & Speece, D. L. (2002). Treatment validity as a unifying construct for identifying learning disabilities. *Learning Disability Quarterly, 25*(1), 33–45.

Interventions INT_Ch10_11.3d 2/2/10 19:56:31
The Charlesworth Group, Wakefield +44(0)1924 369598 - Rev 7.51n/W (Jan 20 2003)

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.

Fuchs, L. S., & Vaughn, S. R. (2005). Response to Intervention as a framework for the identification of learning disabilities. *Forum for Trainers of School Psychologists,* Spring, 12–19.

Gerber, M., & Semmel, M. (1984). Teachers as imperfect tests: Reconceptualizing the referral process. *Educational Psychologist, 19,* 137–148.

Germann, G., & Tindal, G. (1985). Applications of direct and repeated measurement using curriculum based assessment. *Exceptional Children, 51,* 110–121.

Giangreco, M. F., Dennis, R. E., Edelman, S. W., & Chigee, J. C. (1994). Dressing your IEP for the general education climate: Analysis of IEP goals and objectives for students with multiple disabilities. *Remedial and Special Education, 15,* 288–296.

Good, R. H., & 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.

Hintze, J. M., Ryan, A. L., & Stoner, G. (2003). Concurrent validity and diagnostic accuracy of the Dynamic Indicators of Basic Early Literacy Skills and the Comprehensive Test of Phonological Processing. *School Psychology Review, 32,* 541–556.

Hintze, J. M., & Shapiro, E. S. (1997). Curriculum-based measurement and literature-based reading: Is curriculum-based measurement meeting the needs of changing reading curricula. *Journal of School Psychology, 35,* 351–375.

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., & Silberglitt, 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.

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.

Individuals with Disabilities Education Improvement Act of 2004. 34 C.F.R. 2004.

Kame'enui, E. J. (2002). *Final report on analysis of reading assessment instruments for K-3.* Eugene, OR: Institute for Educational Achievement.

Kaminski, R. A., Cummings, K. D., Powell-Smith, K. A., & Good, R. H. (2008). Best practices in using Dynamic Indicators of Basic Early Literacy Skills (DIBELS) for formative assessment and evaluation. In A. Thomas & J. Grimes (Eds.), *Best practices in school psychology V* (pp. 1181–1204). Bethesda, MD: National Association of School Psychologists.

Kaminski, R. A., & Good, R. H. (1996). Toward a technology of assessing basic early literacy skills. *School Psychology Review, 25,* 215–227.

Kaminski, R. A., & Good, R. H. (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.

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., Deno, S. L., & Tindal, G. (1984). Eligibility for learning disabilities services: A direct and repeated measurement approach. *Exceptional Children, 50,* 554–555.

Marston, D., & Magnusson, D. (1985). Implementing curriculum-based measurement in special and regular education settings. *Exceptional Children, 52,* 266–276.

Marston, D., Mirkin, P. K., & Deno, S. L. (1984). Curriculum-based measurement: An alternative to traditional screening, referral and identification of learning disabled students. *Journal of Special Education, 18,* 109–118.

McMaster, K., & Espin, C. (2007). Technical features of curriculum-based measurement in writing. *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. *Journal of Special Education, 41,* 85–120.

National Center on Response to Intervention (NCRTI). (2009). Review of Progress Monitoring Tools. Website of the U.S. Office of Special Education Programs, Progress Monitoring Technical Review Committee: http://www.rti4success.org/chart/progressMonitoring/progressmonitoringtoolschart.htm

Pearson, P. D. (2006). Foreword. In K. S. Goodman (Ed.), *The truth about DIBELS: What it is. What it does* (pp. v–xix). Portsmouth, NH: Heinemann.

Pericola Case, L., Speece, D. L., & Eddy Molloy, D. (2003). The validity of response-to-instruction paradigm to identify reading disabilities: A longitudinal analysis of individual differences and context factors. *School Psychology Review, 32,* 557–582.

Rodden-Nord, K., & Shinn, M. R. (1991). The range of reading skills within regular education classrooms and understanding of special education services for the mildly handicapped. *Journal of Special Education, 24,* 441–453.

Samuels, S. J. (2007). The DIBELS tests: Is speed of barking at print what we mean by reading fluency? *Reading Research Quarterly, 42,* 563–566.

Shapiro, E. S., Angello, L. M., & Eckert, T. L. (2004). Has curriculum-based assessment become a staple of school psychology practice? An update and extension of knowledge, use, and attitudes from 1990 to 2000. *School Psychology Review, 33,* 249–257.

Shin, J., Deno, S. L., & Espin, C. (2000). Technical adequacy of the maze task for curriculum-based measurement of reading growth. *Journal of Special Education, 34,* 164–172.

Interventions INT_Ch10_11.3d 2/2/10 19:56:31
The Charlesworth Group, Wakefield +44(0)1924 369598 - Rev 7.51n/W (Jan 20 2003)

Shinn, M. R. (Ed.). (1989). *Curriculum-based measurement: Assessing special children*. New York: Guilford Press.

Shinn, M. R. (Ed.). (1998). *Advanced applications of curriculum-based measurement*. New York: Guilford Press.

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, Inc.

Shinn, M. R. (2003b). AIMSweb training workbook: Organizing and implementing a benchmark assessment program. Eden Prairie, MN: Edformation, Inc.

Shinn, M. R. (2007). Identifying students at risk, monitoring performance, and determining eligibility within RTI: Research on educational need and benefit from academic intervention. *School Psychology Review, 36,* 601–617.

Shinn, M. R. (2008). 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 V* (pp. 243–262). Bethesda, MD: National Association of School Psychologists.

Shinn, M. R., Good, R. H., Knutson, N., Tilly, W. D., & 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. I., & Parker, C. (1999). Non-categorical special education services with students with severe achievement deficits. In D. J. Reschly, W. D. I. Tilly & J. P. Grimes (Eds.), *Special education in transition: Functional assessment and noncategorical programming* (pp. 81–106). Longmont, CO: Sopris West.

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. Telzrow & 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. (2003). *AIMSweb training workbook administration and scoring of reading maze for use in general outcome measurement*. Eden Prairie, MN: Edformation, Inc.

Shinn, M. R., Tindal, G., & Spira, D. (1987). Special education referrals as an index of teacher tolerance: Are teachers imperfect tests. *Exceptional Children, 54,* 32–40.

Silberglitt, B. (2007). Using AIMSweb early literacy measures to predict successful first-grade readers. Eden Prairie, MN: Edformation, Inc.

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.

Simmons, D. C., Kame'enui, E. J., Good, R. H., Harn, B., Cole, C., & Braun, D. (2002). Building, implementing, and sustaining a beginning reading improvement model: Lessons learned school by school. In M. R. Shinn, H. M. Walker, & G. Stoner (Eds.), *Interventions for academic and behavior problems: Preventive and remedial approaches*. Bethesda, MD: National Association of School Psychologists.

Smith, S. W. (1990). Individualized educational programs (IEPs) in special education: From intent to acquiescence. *Exceptional Children, 57,* 6–14.

Stage, S. A., & Jacobsen, M. D. (2001). Predicting student success on a state-mandated performance-based assessment using oral reading fluency. *School Psychology Review, 30,* 407–419.

Stewart, L. H., & Silberglitt, B. (2008). Best practices in developing academic local norms. In A. Thomas & J. Grimes (Eds.), *Best practices in school psychology V* (pp. 225–242). Bethesda, MD: National Association of School Psychologists.

Thurber, R. S., & Shinn, M. R. (2002). What is measured in mathematics tests? Construct validity of curriculum-based mathematics measures. *School Psychology Review, 31,* 498–513.

Tindal, G., & Germann, G. (1985). Models of direct measurement in the determination of eligibility, monitoring of student progress, and the evaluation of program effects. *British Columbia Journal of Special Education, 9,* 365–382.

Torgesen, J. K. (1998). Catch them before they fall: Identification and assessment to prevent reading failure in young children. *American Educator*, Spring, 32–39.

Torgesen, J. (2004). Lessons learned from the last 20 years of research in the interventions for students who experience difficulty learning to read. In P. McCardle & V. Chhabra (Eds.), *The voice of evidence in reading research* (pp. 225–229). Baltimore: Brookes.

Vaughn, S., & Fuchs, L. S. (2003). Redefining learning disabilities as inadequate response to instruction: The promise and potential problems. *Learning Disabilities Research and Practice, 18,* 137–146.

Walker, H. M., & Severson, H. (1994). Replication of the Systematic Screening for Behavior Disorders (SSBD) procedure for the identification of at risk children. *Journal of Emotional and Behavioral Disorders, 2,* 66–74.

Wikipedia. (n.d.). http://en.wikipedia.org/wiki/

Interventions INT_Ch10_11.3d 2/2/10 19:56:32
The Charlesworth Group, Wakefield +44(0)1924 369598 - Rev 7.51n/W (Jan 20 2003)