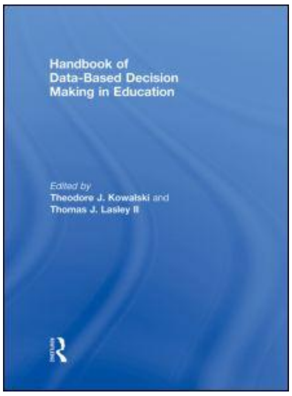

Handbook of Data-Based Decision Making in Education

Theodore J. Kowalski, Thomas J. Lasley

Preparing Educators to Use Curriculum-Based Measurement

Publication details
https://www.routledgehandbooks.com/doi/10.4324/9780203888803.ch9
Robin S. Codding
Published online on: 13 Oct 2008

Handbook of Data-Based Decision Making in Education

Edited by

Theodore J. Kowalski and
Thomas J. Lasley II

First published 2009
by Routledge
270 Madison Ave, New York, NY 10016

Simultaneously published in the UK
by Routledge
2 Park Square, Milton Park, Abingdon, Oxon OX14 4RN

This edition published in the Taylor & Francis e-Library, 2008.

"To purchase your own copy of this or any of Taylor & Francis or Routledge's
collection of thousands of eBooks please go to www.eBookstore.tandf.co.uk."

Routledge is an imprint of the Taylor & Francis Group, an informa business

Library of Congress Cataloging-in-Publication Data
Handbook of data-based decision making in education / Theodore J. Kowalski & Thomas J. Lasley II, editors.
p. cm.
Includes bibliographic references and index.
1. School management and organization–Decision making–Handbooks, manuals, etc. I. Kowalski, Theodor 1943– II. Lasley II, Thomas J. 1947–
LB2805 .H2862 2008
371.2 22

ISBN 0-203-88880-4 Master e-book ISBN

ISBN10: 0–415–96503–9 (hbk)
ISBN10: 0–415–96504–7 (pbk)
ISBN10: 0–203–88880–4 (ebk)

ISBN13: 978–0–415–96503–3 (hbk)
ISBN13: 978–0–415–96504–0 (pbk)
ISBN13: 978–0–203–88880–3 (ebk)

9

Preparing Educators to Use Curriculum-Based Measurement

Robin S. Codding
University of Massachusetts Boston

and

James E. Connell, Jr.
Temple University

Introduction

Academic assessment is multi-faceted and can be used to make various kinds of decisions including: (a) referral, (b) screening, (c) classification, (d) instructional planning, (e) monitoring student progress, and (f) program evaluation (Salvia & Ysseldyke, 2001). Not all forms of assessment can be used to address each of these decisions; therefore, it is critical the measurement tool used matches the pending instructional decision. Curriculum-based Measurement (CBM) was initially created to fill a gap within academic assessment; specifically to generate a tool that was simple, easy to use, and was an accurate indicator of skill proficiency that could be used to monitor student progress (Deno, 1985; Wayman, Wallace, Wiley, Ticha, & Espin, 2007). The most unique feature of CBM compared to other types of assessment is the sensitivity to small changes in student learning, making these tools particularly effective for that purpose. Perhaps one of the most important research findings is that instructional quality and student achievement increase when teachers use CBM to monitor students' performance (Fuchs, Fuchs, Hamlett, & Ferguson, 1992). CBM is also fluency based; that is, these measures are intended to evaluate performance within a brief, pre-determined amount of time (Fuchs & Fuchs, 2004). Evidence has shown that CBM is also useful for making decisions related to screening and instructional planning.

Direct Academic Assessment

CBM is a form of direct academic assessment of basic skills. Direct assessment refers to the direct evaluation of student performance on materials that students are expected to learn; thus, the content of the assessment is derived from curricular objectives (Shapiro, 2004). There are many models subsumed under the umbrella of direct assessment including criterion-referenced curriculum-based assessment

(CBA-CR; Blankenship, 1985), curriculum-based assessment for instructional design (CBA-ID; Gickling & Havertape, 1981), curriculum-based evaluation (CBE; Howell & Nolet, 1999), and CBM (e.g., Deno, 1985; Shinn, 1989). All of these models share the common notion that assessment should test what is taught (Shapiro, 2004). Other essential features of direct academic assessment are that measurements (a) be repeated over time, (b) include valid indicators of critical outcomes of instruction, and (c) inform instructional decisions (Fuchs & Deno, 1994). Research has indicated that these measurements need not be derived from the specific instructional curriculum used in a particular district (Bradley-Klug, Shapiro, Lutz, & DuPaul, 1998; Fuchs & Deno, 1994; Wayman et al., 2007). Consequently, several applications of curriculum-based assessment (CBA) and CBM are available such as Dynamic Indicators of Basic Early Literacy Skills (DIBELS), AIMSwebfi (www.aimsweb.com), and STEEP (www.isteep.com). Readers are encouraged to consult an Office of Special Education Programs sponsored website, entitled National Center on Student Progress Monitoring, which describes and evaluates commercially available CBM/A tools (www.studentprogress.org/chart/chart.asp).

Each form of direct assessment can be categorized as either general outcome or subskill-mastery models (Fuchs & Deno, 1991; Shapiro, 2004). The goal of the general outcome model (GOM) is to assess long-term growth across expected year-end objectives. This is done using standardized measures, for which difficulty level is controlled, that represent "brief, timed samples of performance" (Shapiro, 2004, p. 18). A limitation of this model is that it is not designed to inform specific instructional changes on short-term curricular objectives. However, because GOMs provide indicators of progress toward year-end performance goals, teachers can modify instruction accordingly (Fuchs & Fuchs, 2004). A form of direct assessment that falls into this category is CBM (Shinn, 1989). The content of CBM reflects indicators of student learning often described as "vital signs" that have been demonstrated as valid and reliable through over 20 years of research. That is, each CBM score serves as a gauge of overall competence in the associated academic area (Fuchs & Fuchs, 2004; Shinn, 1989).

Subskill-mastery models address the aforementioned limitations by determining whether students are meeting short-term objectives. Measurements differ as they are based on skill hierarchies and thus a new form of measurement (i.e., probe) is administered in correspondence with each new objective instructed. Furthermore, mastery criteria are identified for each subskill area (Shapiro, 2004). Therefore, these measurements provide more specific details that permit educators to fine-tune delivery of instruction as necessary. CBA-CR (Blankenship, 1985), CBA-ID (Gickling & Havertape, 1981), and CBE (Howell & Nolet, 1999) can be subsumed under the subskill mastery model.

The most effective way to make data-based decisions using direct forms of academic assessment is to combine the use of general outcome and subskill mastery models (e.g., Shapiro, 2004). In this way both long-term and short-term curricular objectives can be monitored. For most students, examining CBM (general outcome) performance is sufficient. However, for students not responding to instruction, conducting a skills analysis consistent with CBA procedures may also be required. Therefore, for the purpose of this chapter we will describe the use of CBM (Deno, 1985) and subskill mastery models of CBA, such as CBA-ID (Gickling & Havertape, 1981) and CBE (Howell & Nolet, 1999), at the elementary level (Grades 1 to 5). We will begin our discussion by describing the following: (a) CBM/A types, (b) administration and scoring procedures, and (c) technical adequacy across reading, mathematics, and writing. Subsequently, instructional decision making will be described for these academic areas and focus on using these tools for screening, progress monitoring, and instructional planning.

Using CBM/A-Reading to Make Instructional Decisions

Types of CBM/A Reading Assessment

There are three general areas of direct assessment in reading: (a) CBM-R, (b) CBM-Maze, and (c) early literacy. These measures address reading fluency, basic comprehension, and beginning literacy skills (e.g., letter-naming fluency, phoneme segmentation), respectively. Because the focus of this chapter is on primary grades (i.e., Grades 1 through 5) the ensuing section will not include discussion of early literacy CBM. Administration of CBM-R requires that one (screening) or three (progress monitoring) passages be provided to students using the appropriate grade-level material (Ardoin et al., 2004; Shinn, 1989). Students read the passage(s) aloud for one minute, while the examiner follows along marking errors. For the purpose of error analysis, examiners might also record the types of mistakes made. Scores are computed by subtracting the number of errors from the total number of words read to yield the number of words read correctly per minute (WRCM; Shinn, 1989). When three passages are administered the median WRCM should be used (see Table 9.1 for administration and scoring procedures).

Test-retest, alternative form, and interobserver agreement indicate strong reliability evidence for CBM-Reading (CBM-R), also described as Oral Reading Fluency (ORF; Marston, 1989). Validity coefficients with published measures of reading range from moderate to strong. A recent review of research (Wayman et al., 2007) confirmed earlier findings by demonstrating that a strong relationship exists between CBM-R and reading proficiency across primary grades. CBM-R was also found to be a better indicator of reading comprehension than other common measures. Additionally, several studies found that CBM-R was strongly correlated with state reading tests. This review also indicated that CBM-R may not be the best choice for beginning or secondary readers. CBM-Maze and early literacy measures may be more useful for middle school and kindergarten students, respectively. Less research has evaluated psychometric properties of subskill mastery models of reading CBA; however, preliminary research on CBA-ID yielded good reliability evidence (Burns, Tucker, Frame, Foley, & Hauser, 2000).

CBM-Maze is a multiple-choice closed passage comprehension measure. Unlike CBM-R, administration of CBM-Maze can be conducted in a group and requires that one, three-minute passage be provided to students. Every seventh word in the passage is replaced with a choice of three words representing the correct choice, a near distracter (i.e., a word of the same type) and a far distracter (i.e., randomly selected).

Table 9.1 General administration and scoring procedures for CBM/A measures.

CBM/A measure	*No. administered*	*Minutes*	*Scores*
CBM-R	3 (1)*	1 Minute	WRCM (Total Words–Errors) WRIM (Words Incorrect)
CBM-Maze	1	3 Minutes	Correct Choices
CBA-M (Single Skill)	1	2–5 Minutes	CD (Digits Counted as Correct) DCPM (Total Digits–Errors/2) DIPM (Digits Incorrect/2)
CBM-M (Multiple Skill)	3	2–5 Minutes	CD (Digits Counted as Correct) DCPM (Total Digits–Errors/2) DIPM (Digits Incorrect/2)
CBM-WW	1	3 Minutes	WW (Total Words Written)
CBM-CWS	1	3 Minutes	CWS (Correct Writing Sequences)
CBM-WSC	1	3 Minutes	WSC (Words Spelled Correct)

* (1) passage may be used for universal screening (Ardoin et al., 2004). CBA-M = Curriculum-Based Assessment Mathematics; CBM-M = Curriculum-Based Measurement Mathematics; WRCM = words read correctly per minute; WRIM = words read incorrectly per minute; DCPM = digits correct per minute; DIPM = digits incorrect per minute. These are basic guidelines (Gansle et al., 2006; Shinn, 1989; Shinn & Shinn, 2002), although depending on application used these procedures may vary slightly.

Students are instructed to circle the correct answer and scores are computed by counting the number of correct word choices (Shinn & Shinn, 2002). Findings for CBM-Maze indicate that this measure has adequate reliability and validity across Grades 2 through 8 and growth rates were more consistent than CBM-R (Wayman et al., 2007).

Screening Using CBM: Reading Benchmark Assessments

Within the framework of multi-tiered models of instruction, CBM-R and CBM-Maze can be used as indicators of all students' reading proficiency. In this way, CBM is used to screen for students at risk; that is, students who may need academic interventions in addition to instruction in the grade-level curriculum. Universal screening occurs three times yearly (fall, winter, and spring), and is often referred to as benchmark assessments. Theoretically, it is anticipated that approximately 20% of students will need support in addition to the core instruction, presuming that instruction is conducted using scientifically validated curricula (Batchse et al., 2006).

Schoolwide data at each grade level can be examined according to percentiles and averages obtained from national normative samples, such as provided by AIMSwebfi, or can be examined at the local level (Shinn, 1989). Generally, students falling below the 25^{th} percentile are considered to be at risk (Fuchs, 2003) and those students performing below the 16^{th} percentile have been described as likely to have a valid special education referral (VanDerHeyden, Witt, & Naquin, 2003). Two basic decisions are made using CBM-R or CBM-Maze data at this level. First, students above the 25^{th}

percentile are confirmed to be responsive to core instruction (Fuchs, 2003). Second, students below the 25th percentile are considered to be potentially at risk. Once the at-risk students are identified, then another set of decisions needs to be made; will at-risk students: (a) be monitored weekly or bi-weekly for four to eight weeks (Hintze, 2007; Shapiro, 2004), or (b) receive additional instructional support? It is possible that some at-risk students may be monitored while others require instructional interventions. For example, a school building-level team may decide to provide additional reading interventions to students below the 16th percentile (option b; VanDerHeyden et al., 2003) while monitoring all at-risk students below the 25th percentile (option a). Educators might also combine the CBM results with other data collected such as previous performance on state assessments or other standardized measures (e.g., Iowa Tests of Basic Skills; Hoover, Hieronymus, Frisbie, & Dunbar, 1993) to determine which students should receive additional instructional support.

Although initial decisions are made after the fall CBM-R or CBM-Maze data are collected, screening in the winter and spring allows educators to determine whether students previously identified as responding to core instruction continue to perform well. That is, some students who performed well during the fall screening may read in the at-risk range according to the winter benchmarks. Therefore, the same sets of decisions made after examining the fall data should also be made after CBM data are collected in the winter and spring.

Progress Monitoring and Instructional Planning Using CBM/A-Reading

Monitoring At-Risk Students Not Receiving Additional Academic Support For students that are determined to be at risk following the initial benchmark assessment, reading progress in response to the core curriculum should be monitored weekly or bi-weekly using the same CBM procedures described earlier but using alternate reading passages (Hintze, 2007). That is, a new passage is administered each week. Student performance should be graphed, which could be done by using Microsoft Excelfi or other data management systems (e.g., isteep, AIMSwebfi, Monitoring Basic Skills Progress; Fuchs, Hamlett, & Fuchs, 1997), and evaluated for improvement. Improvement is determined by looking at the level of performance (e.g., did performance reach or exceed the criterion?) as well as the slope of progress (i.e., is performance increasing?). Measurement of performance level *along with growth*, as described here, uses a dual discrepancy approach to decision making (Fuchs, 2003).

The level of performance can be determined by identifying whether the student reached the 25th percentile at the end of eight weeks using local or national norms or examining whether the student has reached a criterion-referenced benchmark, such as reading 40 words correctly per minute (WRCM) in first-grade material (Fuchs, 2003; Good, Simmons, & Kame'enui, 2001). In order to evaluate growth, a realistic but ambitious goal needs to be established and therefore it is necessary to know how much growth should be expected (Fuchs, Fuchs, Hamlett, Walz, & Germann, 1993). If these data have not been collected by the local school district, then data from Fuchs et al. (1993) can be used (Table 9.2 provides criteria and growth data for CBM across academic areas). For example, a third-grade student reading 60 WRCM on the third

Table 9.2 Common benchmarks and growth rates.

Measure	*Grade*	*Benchmark*	*Growth*[a]
CBM-R	1	60 WRCM[b]	2.00–3.00 Words/Week
	2	60 WRCM	1.50–2.00 Words/Week
	3	100 WRCM	1.00–1.50 Words/Week
	4	100 WRCM	0.85–1.10 Words/Week
	5	100 WRCM	0.50–0.80 Words/Week
CBM-Maze	1–5	—	0.39, 0.84 Replacements
CBA/M-M	1–3	20 DCPM[c]	0.30 Digits/Week
	4	40 DCPM	0.70 Digits/Week
	5	40 DCPM	0.75 Digits/Week
CBM-W	1	15 WW[d]	—
	2	28 WW	—
	3	37 WW	—
	4	41 WW	—
	5	49 WW	—

Note: Benchmark numbers for reading and mathematics represent the mastery-level criterion. CBA/M-M growth rates are based on multiple-skill probes. These criteria are from the following sources: [a]Fuchs et al. (1993), [b]Fuchs & Deno (1982), [c]Deno & Mirkin (1977), and [d]Mirkin et al. (1981).

grade fall CBM benchmark is performing below the 25th percentile (Fuchs, 2003; Fuchs & Deno, 1982). According to Fuchs et al. (1993), third-graders should gain between 1 and 1.5 words read (aloud) per week. In order to establish the student's goal, the average learning rate (i.e., 1.0 or 1.5 WRCM) is multiplied by the number of weeks the student will be monitored (eight weeks) and this number is added to the current reading score ([1.0 WRCM × 8 weeks] + 60 WRCM = 68 WRCM). In this example, it would be expected that the student read 68 WRCM by the end of eight weeks. This goal can be plotted on a graph and becomes the aim line for the student (see Figure 9.1, Panel 1). These same procedures can be followed using CBM-Maze. Unlike CBM-R, expected growth may not be different depending on grade level. Preliminary data suggest that targets for weekly improvement may be .39 and .84 words correct for realistic and ambitious goals, respectively (Fuchs et al., 1993). For example, if a third-grade student obtained a score of 7 at the fall benchmark, then the ambitious goal at the end of eight weeks may be 14 ([0.84 × 8 weeks] + 7 = 13.72).

A goal can also be established by expecting a student to reach the 25th percentile for the fall or winter benchmark criterion by the end of the monitoring period (Shapiro, 2004). Continuing with this example, the student would be expected to read between 65 and 70 WRCM (Hasbrouck & Tindal, 1992). This might be particularly useful when monitoring at-risk students not yet receiving supplemental reading support. If a student reaches the expected goal by the end of the monitoring period or is making adequate progress toward this goal, then the student is responding to the core curriculum. If, however, a student is not reaching the expected level of performance and is not progressing, additional reading support may be necessary.

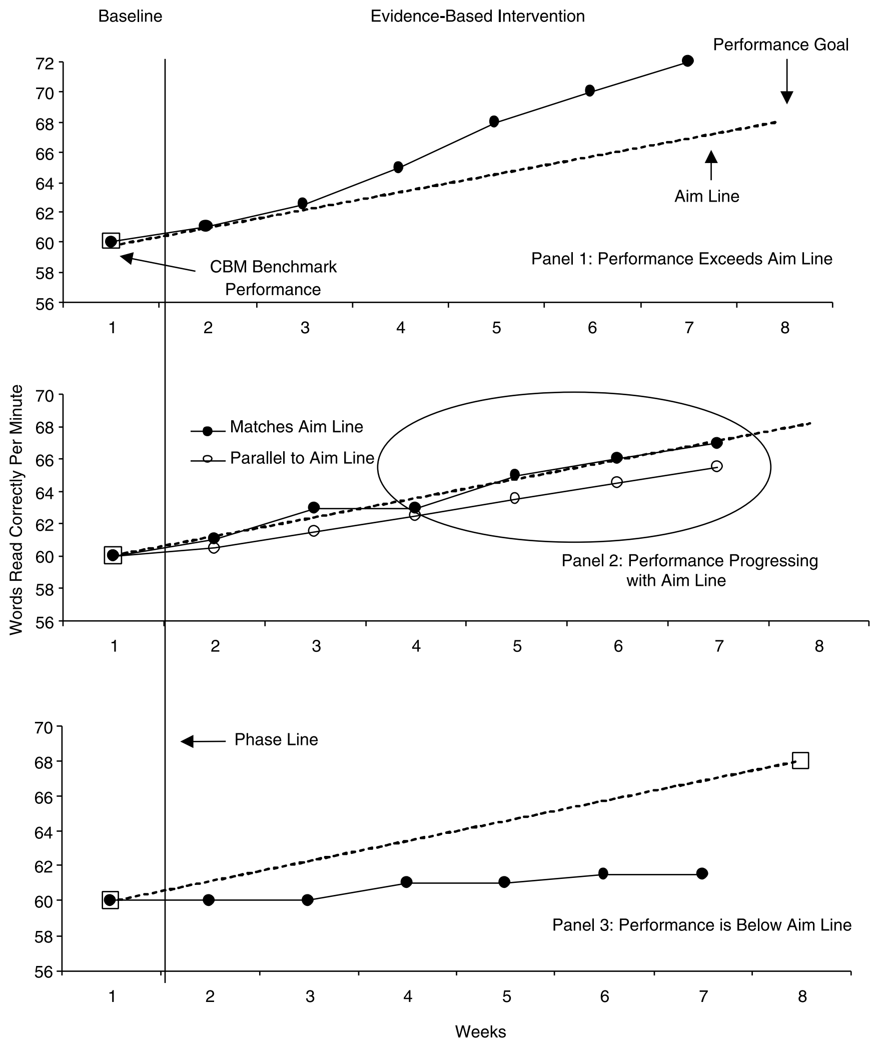

Figure 9.1 Instructional decisions when CBM performance exceeds, meets, or is below the aim line (goal).

Monitoring At-Risk Students Receiving Additional Academic Support When students are not performing as expected in the core curriculum according to the earlier described decision rules, two levels of academic support may be identified. The first may be within small groups using empirically supported supplemental reading interventions. The second may be individualized reading support, often provided when response to small group interventions is unsuccessful (Batsche et al., 2006). For these students it is important to have more information regarding specific reading weaknesses so that appropriate intervention planning can occur. This can be done by analyzing the CBM data to determine what kind of reading problem appears to be evident or by using a subskill mastery model of CBA (e.g., CBE).

A first step may be to determine whether students have reading fluency, decoding, or comprehension problems. This can be done by examining the WRCM and the errors made. Specifically, if a student is making two or fewer errors when reading grade-level material but WRCM are falling below the 25th percentile, s/he may be exhibiting reading fluency problems (Daly, Witt, Martens, & Dool, 1997; Gravois & Gickling, 2002). For students who exhibit low WRCM and are reading with many errors, decoding may be problematic. Finally, in order to identify comprehension problems it might be useful to examine performance on CBM-Maze or generate five to eight open-ended questions (literal and inferential) to be answered following administration of a CBM-R probe (Shapiro, 2004). Students with comprehension difficulties may exhibit low fluency on CBM-R and low scores on CBM-Maze. Corresponding small group interventions should match student areas of need.

A second step may be to conduct a survey-level assessment (SLA) to determine the grade-level material that provides the appropriate instructional match (Howell & Nolet, 1999; Shapiro, 2004). In this process material that was mastered and that which is too difficult (frustrational) for students to read is identified. In order to conduct the SLA, examiners need to obtain CBA passages from various grade levels. The SLA begins with material (i.e., CBA probe) that corresponds with a student's grade of record. Consistent with CBA procedures, three passages are presented and the median score is computed. Additionally, a miscue analysis should be conducted during which types of errors, such as omissions, substitutions, transpositions, and hesitations (three-second pause), are recorded. Next, the median WRCM and errors are compared to placement criteria. These criteria could be generated from local or national normative data such that below the 25th percentile would represent the frustrational level, performance between the 25th and 50th percentile would represent the instructional level, and above the 75th percentile would represent mastery (Shapiro, 2004). Fuchs and Deno (1982) also provide placement criteria. If a student's performance falls in the frustrational range on grade-level material, then passages from the previous grade should be administered and compared to the placement criteria. This process continues until the appropriate instructional, mastery, and frustrational levels are identified. SLA permits the analysis of specific skill development so that individualized reading interventions can be generated in which students have the opportunity to read material at their instructional level using strategies that address individual error patterns.

Short-term goals for these students should be established as described above and evaluated to determine if the type and the intensity (e.g., number of minutes or sessions per week) of support is leading to reading improvements. Growth is measured by comparing the actual student performance with the aim line that was generated from the goal. Instructional decisions can be made accordingly using simple decision rules (Fuchs & Fuchs, 2007; Hintze, 2007). First, when looking at these data visually, the most recent four data points should be isolated to determine whether they are going up (increasing), going down (decreasing) or are variable. Second, it is determined whether performance: (a) matches the aim line, (b) exceeds the aim line, or (c) is below the aim line and unlikely to reach goal. If performance *matches* the aim line (see Figure 9.1, Panel 2) then the current intervention procedures should be continued until the goal is reached (Hintze, 2007). Once the goal has been achieved and if the goal reflects an adequate level of progress (e.g., 30th percentile of winter CBM benchmark in grade-level material), then supplemental reading support may be reduced. Similarly, if progress is parallel to the aim line, a decision may be made to wait and see if performance will improve at the expected rates. If the most recent four data points *exceed* the aim line (see Figure 9.1, Panel 1), the goal should be increased. Finally, if these data do not illustrate progression toward the established goal (see Figure 9.1, Panel 3), then a change in the intervention procedures may be warranted (Fuchs & Fuchs, 2007). Once the intervention is changed, a solid vertical line (phase line) would be added to the graph so that performance under both interventions can be evaluated. In each of these circumstances progress should continue to be monitored (at least weekly) and evaluated.

Using CBM/A-Mathematics to Make Instructional Decisions

Types of CBM-Mathematics Assessment

There are three general areas that comprise CBM-M: (a) computation/operations, (b) applications/problem solving, and (c) early numeracy. The most frequently researched form of CBM-M assessment is computation, including single skill and multiple skill operations demonstrating concurrent validity (Connell, 2005; Foegen, Jiban, & Deno, 2007; Skiba, Magnuson, Martson, & Erickson, 1986), reliability (Fuchs, Fuchs, & Hamlett, 1990; Tindal, German, & Deno, 1983), and progress monitoring of performance indicators (Fuchs, Fuchs, Hamlett, & Stecker, 1991). Applications-type assessments (e.g., word problems, algebra, data analysis, measurement, geometry, and patterns) have less supporting data, though the research has found that these measures are reliable (Fuchs et al., 1994) and valid indicators of student performance on standardized assessments (Connell, 2005; Fuchs et al., 1994). Additionally, the applications assessments have high face validity with teachers (Connell, 2005). Correspondence with state assessments range from moderate to strong for computation and applications/problem-solving measures across primary and middle school grades (Shapiro, Keller, Lutz, Santoro, & Hintze, 2006). Finally, and important to note for the remaining discussion, research demonstrates that the constructs representing applications and computation problems, though separate, *are related* (Thurber, Shinn, & Smolkowski, 2002). To date, no studies were identified that investigated early numeracy or application problems in a response-to-intervention framework, where benchmark scores and growth rates were provided. Therefore, the remainder of this section will be limited to computation assessments.

Screening Using CBM: Computation Benchmark Assessments

Within the GOM, the *primary* CBM-M datum is described as the "performance indicator" of overall math proficiency. The performance indicator is a measure of individual student performance on all skills targeted for instruction across the year (Fuchs et al., 1990). School districts across the nation are presently developing "benchmark assessments." And, whether developed in schools by district personnel or outsourced to test development agencies (e.g., Educational Testing Service), the purpose of these benchmark assessments is to acquire a performance indicator in a specific content area (e.g., mathematics, literacy) for each student assessed. It is expected that with this performance indicator, teachers will differentiate instruction to meet each student's instructional needs.

Like CBM-R, universal screening in mathematics occurs three times per year, fall, winter and spring, and can be group or individually administered (VanDerHeyden et al., 2003; Witt, Daly, & Noell, 2000). As stated above, computation assessments are used because they are reliable and valid measures of overall mathematics performance, are easy to administer, are brief, can be used diagnostically, and can be used to monitor progress. The administration procedure is as follows. First, assessments are developed that represent, in equal number, the types of skills students are expected to learn across the year. There are typically 25 to 30 problems on a page. For example, a third-grade performance indicator assessment might include five multi-digit addition with regrouping and five multi-digit addition without regrouping, five multi-digit subtraction with borrowing and five multi-digit subtraction without borrowing, and five multiplication facts to 9, totaling 25 computation problems (Hosp, Hosp, & Howell, 2007). If group administered, assessments are placed face down on each student's desk. Students are instructed that the sheet contains several types of math facts (e.g., addition, subtraction, and multiplication). Next, students are told they have two minutes (three to five minutes for more complex skills) to complete as many problems as they can, and attend to the operation. Then, students are told to begin the assessment and the timer is started. After two minutes, students are told to stop and the assessments are collected. Assessments are scored by counting the number of correct digits per two-minute assessment. Correct digits (CD) are determined by counting the number of CD in the answer to the problem (attending to place value). Each correct digit is given one point. Points are tallied and the total number of points (i.e., CD) is the student's score for that assessment. Once the data are collected and analyzed, school teams are able to evaluate if the core curriculum is effective for most students by comparing the assessment scores to the local or national norms (Burns, VanDerHeyden, & Jiban, 2005; Deno & Mirkin, 1977).

Performance indicators are effective screening tools because they allow for individual basic skills to be assessed. In other words, by assessing the skills students are expected to learn across the year, educators can determine which skills have yet to be acquired for which students. Thus the performance indicator can be used diagnostically. For example, if a student's assessment data illustrate that multi-digit addition without regrouping has been established (i.e., all digits attempted are correct), but multi-digit addition with regrouping has not been acquired (i.e., no correct digits), then educators can determine, by making data-based decisions, if more intensive intervention is required on the missing skill(s).

Progress Monitoring and Instructional Planning Using CBM/A-M

Monitoring At-Risk Students Not Receiving Additional Academic Support Progress monitoring is recommended for students identified as at risk (below the 25th percentile) following universal screening. Progress monitoring includes weekly or bi-weekly (Hosp et al., 2007) re-administration of the performance indicator assessment used in universal screening for up to eight weeks. Graphing the data by pencil or with a computer-based graphing application allows for visual analysis of performance over time (i.e., slope). Following the dual-discrepancy model, level and rate of growth are evaluated. If the at-risk student reaches or exceeds the 25^{th} percentile (i.e., level) after the monitoring period, then the student is no longer considered at risk. If the rate of growth matches or exceeds those performing above the 25^{th} percentile, then progress monitoring continues until the student reaches the expected level of performance (i.e., local norm or criterion-referenced benchmark score). If the level of performance does not reach the local or national norms by the end of the eight weeks and the slope is less than the local or national growth rates, then intervention is needed.

Monitoring At-Risk Students Receiving Additional Academic Support Skills-based measures (SBM) are assessments used to screen and monitor one mathematics basic skill (i.e., addition, subtraction, multiplication, and division). Thus, the *secondary* datum is the CBA skills analysis (Fuchs et al., 1990). A skills analysis is a process whereby performance on any one skill is analyzed to determine skill proficiency, monitor progress, and guide instruction. Following the example above, all classroom students are reassessed on multi-digit addition with regrouping to determine if the skill deficit is a result of ineffective curriculum and/or instruction. If the SBM universal screening data illustrate that only a few students have not acquired this skill, then intervention and progress monitoring is initiated for these students in small group format using evidence-based interventions (Ardoin, Witt, Connell, & Koenig, 2005). For each student, the SBM universal screening performance level serves as the baseline data point from which projected slope is calculated. For example, if John's SBM assessment score was 15 CD/two minutes, then 15 would be the baseline data point. A realistic goal across eight weeks of intervention would be an increase of .3 CD per week (Fuchs et al., 1993) from 15 CD to 17.4 CD per two minute assessment. An ambitious goal would be an increase of .5 CD per week from 15 to 19 CD per week after eight weeks of intervention. The projected realistic and ambitious slope data would be presented on a graph such as that illustrated in Figure 9.1, Panel 1. It is recommended that progress monitoring occur at least weekly and preferably, twice weekly (Hosp et al., 2007). Finally, the student's performance level on the SBM (multi-digit addition with regrouping) is recorded on the same graph as the projected slope for direct visual comparison.

Using CBM-Writing to Make Instructional Decisions

Types of CBM-Writing Assessments

CBM-writing (CBM-W) has remained consistent over the years in terms of the administration procedures for elementary and secondary students (Deno, Mirkin, & Marston 1980; Shinn, 1989). The assessment tool consists of a story starter (also called an incomplete stimulus prompt, or sentence fragment) as described by

Gansle, VanDerHeyden, Noell, Resetar, and Williams (2006). First, students are given a story starter on lined paper and instructed to think about what they will write for the next minute. After one minute has passed, students are instructed to construct a story from the opening sentence provided. Finally, students are told they have three minutes[1] to write, the timer is started and when the timer stops, the writing samples are collected. This procedure can be administered individually for progress monitoring or in group format for screening.

Scoring CBM-W assessments has been less consistent through the years as researchers have evaluated procedures that are reliable and valid indexes of student writing performance across and within grade level(s). The results of those studies indicate that as writing instruction and expectations increase as the grade level increases, more complex scoring methods are then required. Therefore, a brief description of the scoring methods will follow as will a selected sample of reliability and validity studies. For a thorough review of the CBM-W research, see McMaster and Espin (2007).

A number of scoring procedures have been evaluated for the three-minute writing sample. Fluency-based CBM-W measures (i.e., a three-minute assessment as described above) have found total words written (WW), words spelled correctly (WSC), and correct word sequences (CWS) to have adequate reliability coefficients with criterion measures such as the Test of Written Language (Hammill & Larsen, 1978) and the Stanford Achievement Test (Deno, Marston, & Mirkin, 1982; Gansle, Noell, VanDerHeyden, Naquin, & Slider, 2002; Madden, Gardner, Rudman, Karlsen, & Merwin, 1978). Other scoring procedures such as correct letter sequences, correct minus incorrect word sequences, and percentage of legible words also have adequate reliability coefficients, but are less researched. Much of the extant literature identifies WW and WSC (for primary grades) and WSC and CWS (for secondary grades) as reliable and valid indicators of overall writing performance. Finally, when using the Stanford-9 measures of written expression as the criterion measure, CBM-W scoring procedures including WW, WSC and CWS were found to be valid and reliable indicators of student writing performance. The Six Trait model (used in many schools throughout the United States) was not an adequate predictor of written expression when the criterion measure was the Stanford-9 written language subtests (Gansle et al., 2006). Finally, teachers' holistic ratings of students' written expression indicate less support for WW, and more support for WSC and CWS (Parker, Tindal, & Hasbrouck, 1991). However, as a universal screener, teams will need to consider the time it takes to score WSC and CWS (see Hosp et al. (2007) for explicit scoring instructions).

Screening Using CBM: Writing Benchmark Assessments

As a universal screening tool, Parker et al. (1991, Study 1) found WW, WSC, and CWS to be suitable scoring procedures for second- through fourth-grade students, though WSC was considered the most viable scoring procedure across all the grades and has moderate correlations with teachers' holistic ratings of student writing performance. Parker et al. (1991, Study 2) found that CWS was a sufficient predictor of writing when correlated with holistic ratings of student writing performance for students in Grades 6 through 11.

Little research is available identifying percentiles for *at risk* and *in need of additional support*. AIMSWebﬁ and Hosp et al. (2007) do provide percentiles for CWS for Grades 1 through 8. Teams will need to decide if students below the 25th percentile will receive additional support, be monitored across a specified amount of time (e.g., eight weeks), or if a more diagnostic process is required. For example, school teams could use the 25th percentile and below on WW as the lower limit for the school/district at-risk category. Students' written expression assessments below the 25th percentile could then undergo a more diagnostic scoring procedure such as WSC and CWS. Next, students falling below the 25th percentile when using WSC and CWS scoring procedures would require additional instructional support, and those above the 25th percentile of CWS fall into the at-risk group for progress monitoring.

Progress Monitoring and Instructional Planning Using CBM-W

Monitoring At-Risk Students Not Receiving Additional Academic Support Following the example provided above, students not receiving additional academic support are monitored for progress over an eight- to ten-week period for growth. Sensitivity to growth has been demonstrated using WW and WSC measures in first through sixth grade. One to two writing samples per week over the next eight to ten weeks (Marston, 1982; Shinn, 1981) allow enough time for adequate instruction to occur and related performance to reflect growth. The goal is to establish that students will respond to instruction, as demonstrated by an increase in performance across the time monitored from at risk to above the 25th percentile in CWS. If students do not respond to instruction, additional academic support is needed.

Monitoring At-Risk Students Receiving Additional Academic Support When students do not respond to the general education curriculum in writing, then more intensive writing instruction is required. Again, progress monitoring should occur one to two times per week across the length of the intervention. Secondary and tertiary interventions can occur in small groups or as individual instruction. As stated above, WW and WSC measures have demonstrated sensitivity to instruction (Marston, 1982; Shinn, 1981) and therefore should be used as progress monitoring measures. However, CWS is the more complex measure and therefore the goal is to move the student from below the 25th percentile on this measure to above the 25th percentile (whether using local or national norms). Eight to ten weeks of intervention can include as little as 20 minutes and up to one hour of additional instruction per day, depending on the student support needed (e.g., some support, intensive support).

Incorporating CBM/A into IEP Objectives

CBM/A information may also be incorporated into the Individualized Educational Program (IEP) for students identified as having specific reading, mathematics,

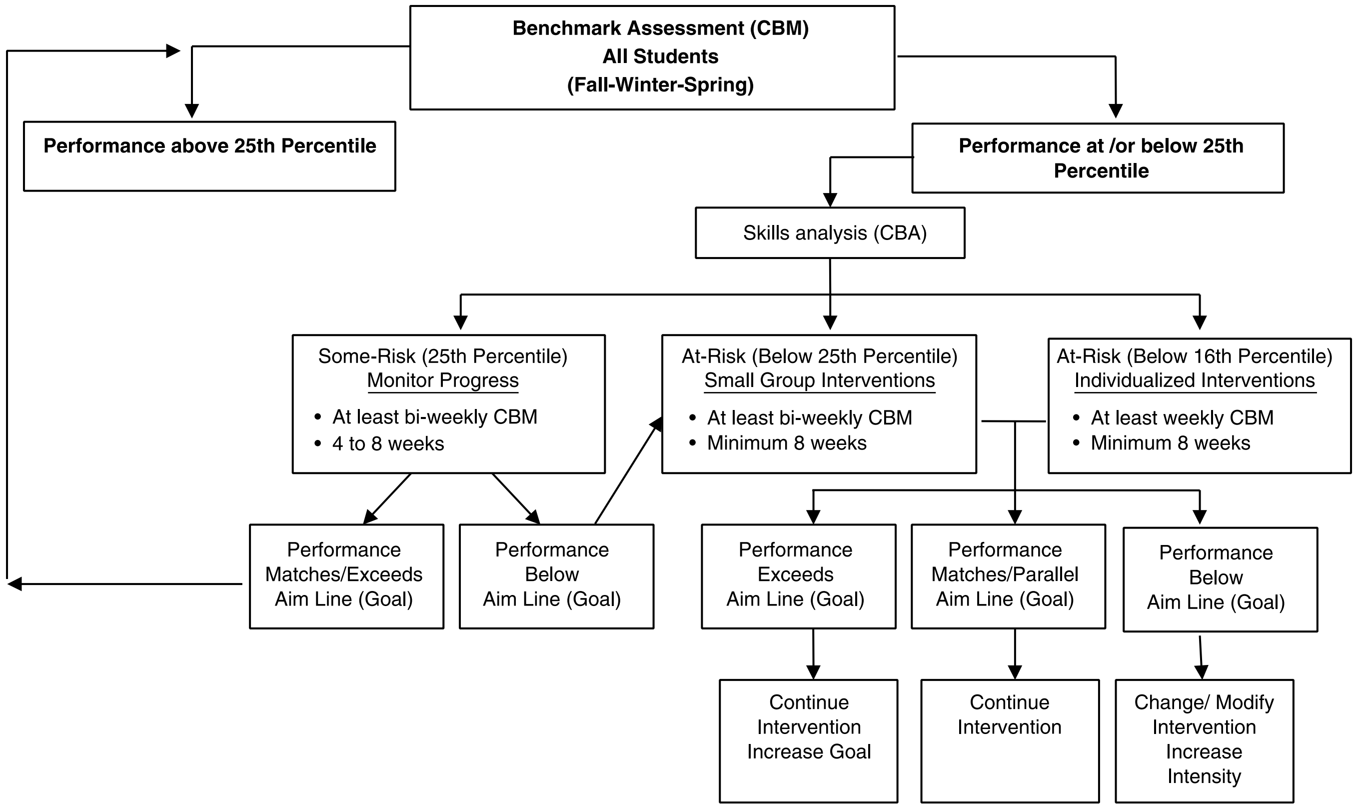

Figure 9.2 Instructional decision making for screening and progress monitoring.

or writing disabilities, or for those students in need of intensive individualized academic support after failing to respond to less intensive evidence-based treatments. CBM/A provides academic goals in the form of observable, measurable, and objective behaviors and criteria (Codding, Skowron, & Pace, 2005; Shinn, 1989). CBM data (fall benchmark), or as determined by the CBA skills analysis (i.e., SLA or SBM), would be listed under the current performance section of the IEP. Using CBM-R as an example, the IEP may read: Given three randomly selected third-grade reading passages from <insert reference to an appropriate curriculum or application>, Molly read aloud a median of 60 WRCM with two errors. The annual goal would reflect the year-end expected level of progress, such as the 50th percentile on the spring CBM benchmark assessment. The objectives to achieve the annual goal would be generated using the procedures described earlier (incorporating average learning rates according to local norms or Fuchs et al., 1993). For example, the first objective may be written as: In 12 weeks, when given three randomly selected third-grade reading passages from <insert reference to an appropriate curriculum or application>, Molly will read aloud a median of 72 WRCM minute with two errors. This goal was determined by multiplying the average learning rate (1.0) by the number of weeks for the objective (12) and adding the product to the fall CBM benchmark (60 WRCM).

Summary

Increased accountability requirements and adoption of response-to-intervention frameworks require that educators utilize measures that are sensitive to small gains in student performance, which also correspond with and inform short-term and year-end curricular objectives. CBM and CBA are particularly suitable for this purpose and the school environment given that these measures are time and resource efficient. The diagram in Figure 9.2 provides an overview of the ways in which CBM and CBA can be used for screening, instructional planning, and monitoring progress.

Note

1. Writing time ranged from one to five minutes in Deno et al. (1980) with the strongest validity coefficients for writing times of three to five minutes with Test of Written Language.

References

Ardoin, S. J., Witt, J. C., Connell, J. E., & Koenig, J. L. (2005). Application of a three-tiered response to intervention model for instructional planning, decision making and the identification of children in need of services. *Journal of Psychoeducational Assessment*, *23*, 362–380.

Ardoin, S. P., Witt, J. C., Suldo, S. M., Connell, J. E., Koenig, J. L., Resetar, J. L., et al. (2004). Examining the incremental benefits of administering a maze and three versus one curriculum-based measurement reading probes when conducting universal screening. *School Psychology Review*, *33*, 218–233.

Batchse, G., Elliott, J., Graden, J. L., Grimes, J., Kovaleski, J. F., Prasse, D., et al. (2006). *Response to intervention: Policy considerations and implementation.* Alexandria, VA: National Association of State Directors of Special Education, Inc.

Blankenship, C. S. (1985). Using curriculum-based assessment data to make instructional management decisions. *Exceptional Children*, *42*, 233–238.

Bradley-Klug, K. L., Shapiro, E. S., Lutz, G. J., & DuPaul, G. J. (1998). Evaluation of oral reading rate as a curriculum-based measure within literature-based curriculum. *Journal of School Psychology*, *36*, 183–197.

Burns, M. K., VanDerHeyden, A. M., & Jiban, C. L. (2005). Assessing the instructional level for mathematics: A comparison of methods. *School Psychology Review*, *35*, 401–418.

Burns, M. K., Tucker, J. A., Frame, J., Foley, S., & Hauser, A. (2000). Interscorer, alternate-form, internal consistency, and test-retest reliability of Gickling's model of curriculum-based assessment for reading. *Journal of Psychoeducational Assessment*, *18*, 353–360.

Codding, R. S., Skowron, J., & Pace, G. M. (2005). Making data-based decisions: Training teachers to use assessment data to create instructional objectives. *Behavioral Interventions*, *20*, 165–176.

Connell, J. E. (2005). Constructing a math applications, curriculum-based assessment: An analysis of the relationship between applications problems, computation problems and criterion-referenced assessments (Doctoral dissertation, Louisiana State University and Agriculture & Mechanical College, 2005). *Dissertation Abstracts International*, 3933.

Daly, E. J., Witt, J. C., Martens, B. K., & Dool, E. J. (1997). A model for conducting a functional analysis of academic performance problems. *School Psychology Review*, *26*, 554–574.

Deno, S. L. (1985). Curriculum-based measurement: The emerging alternative. *Exceptional Children*, *52*, 219–232.

Deno, S. L., & Mirkin, P. K. (1977). *Data-based program modification: A manual*. Reston, VA: Council for Exceptional Children.

Deno, S. L., Martson, D., & Mirkin, P. (1982). Valid measurement procedures for continuous evaluation of written expression. *Exceptional Children Special Education and Pediatrics: A New Relationship*, *48*, 368–371.

Deno, S. L., Mirkin, P. K., & Marston, D. (1980). *Relationships among simple measures of written expression and performance on standardized achievement tests* (Research Report No. 22). Minneapolis: Institute for Research in Learning Disabilities, University of Minnesota.

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

Fuchs, L. S. (2003). Assessing intervention responsiveness: Conceptual and technical issues. *Learning Disabilities Research & Practice*, *18*, 172–186.

Fuchs, L. S., & Deno, S. L. (1982). *Developing goals and objectives for educational programs* [Teaching guide]. Minneapolis: Institute for Research in Learning Disabilities, University of Minnesota.

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. (1994). Must instructionally useful performance assessment be based in the curriculum? *Exceptional Children*, *61*, 15–21.

Fuchs, L. S., & Fuchs, D. (2004). Determining adequate yearly progress from kindergarten through grade 6 with curriculum-based measurement. *Assessment for Effective Intervention*, *29*, 25–37.

Fuchs, L. S., & Fuchs, D. (2007). *Using curriculum-based measurement for progress monitoring in math*. Retrieved November 19, 2007, from http://www.studentprogress.org/summer_institute/2007/math/Mathmanual_10–4.pdf

Fuchs, L. S., Fuchs, D., & Hamlett, C. L. (1990). The role of skills analysis in curriculum-based measurement in math. *School Psychology Review*, *19*, 6–22.

Fuchs, L. S., Hamlett, C. L., & Fuchs, D. (1997). *Monitoring basic skills progress: Basic reading* (2nd ed.) [Computer software manual]. Austin, TX: PRO-ED.

Fuchs, L. S., Fuchs, D., Hamlett, C. L., & Ferguson, C. (1992). Effects of expert system consultation within curriculum-based measurement, using a reading maze task. *Exceptional Children*, *58*, 436–450.

Fuchs, L. S., Fuchs, D., Hamlett, C. L., & Stecker, P. M. (1991). Effects of curriculum-based measurement and consultation on teacher planning and student achievement in mathematics operations. *American Educational Research Journal*, *28*, 617–641.

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

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

Gansle, K. A., Noell, G. H., VanDerHeyden, A. M., Naquin, G. M., & Slider, N. J. (2002). Moving beyond total words written: The reliability, criterion validity, and time cost of alternate measures for curriculum-based measurement in writing. *School Psychology Review*, *31*, 477–497.

Gansle, K. A., VanDerHeyden, A. M., Noell, G. H., Resetar, J. L., & Williams, K. L. (2006). The technical adequacy of curriculum-based and rating-based measures of written expression for elementary school students. *School Psychology Review*, *35*, 435–450.

Gickling E. E., & Havertape, J. (1981). *Curriculum-based assessment*. Minneapolis, MN: National School Psychology Inservice Training Network.

Good, R. H., Simmons, D. C., & Kame'enui, E. J. (2001). The importance and decision-making utility of a continuum of fluency-based indicators of foundational reading skill for third grade high-stakes outcomes. *Scientific Studies of Reading*, *5*, 257–288.

Gravois, T. A., & Gickling, E. E. (2002). Best practices in curriculum-based assessment. In A. Thomas & J. Grimes (Eds.), *Best practices in school psychology IV* (pp. 885–898). Washington, DC: National Association of School Psychologists.

Hammill, D. D., & Larsen, S. C. (1978). *Test of written language.* Austin, TX: PRO-ED.

Hasbrouck, J. E., & Tindal, G. (1992). Curriculum-based oral reading fluency norms for students in grades 2 through 5. *Teaching Exceptional Children*, *24*, 41–44.

Hintze, J. (2007). *Using student progress monitoring in a response to intervention model.* Retrieved November 19, 2007, from http://www.studentprogress.org/library/Webinars.asp#RTI

Hoover, H. D., Hieronymus, A. N., Frisbie, D. A., & Dunbar, S. B. (1993). *Iowa test of basic skills.* Itasca, IL: Riverside Publishing.

Hosp, M. K., Hosp, J. L., & Howell, K. W. (2007). *The ABCs of CBM: A practical guide to curriculum-based measurement.* New York: Guilford Press.

Howell, K. W., & Nolet, V. (1999). *Curriculum-based evaluation: Teaching and decision making* (3rd ed.). Belmont, CA: Wadsworth.

Madden, R., Gardner, E. F., Rudman, H. C., Karlsen, B., & Merwin, J. C. (1978). *Stanford achievement test.* New York: Harcourt Brace Jovanovich.

Marston, D. B. (1982). *The technical adequacy of direct, repeated measurement of academic skills in low-achieving elementary students.* Unpublished doctoral dissertation, University of Minnesota, Minneapolis.

Marston, D. B. (1989). A curriculum-based measurement approach to assessing academic performance: What it is and why do it. In M. R. Shinn (Ed.), *Curriculum-based measurement: Assessing special children* (pp. 18–78). New York: Guilford Press.

McMaster, K., & Espin, C. (2007). Technical features of curriculum-based measurement in writing. *Journal of Special Education*, *41*, 68–84.

Mirkin, P. D., Deno, S. L., Fuchs, L. S., Wesson, C., Tidal, G., Marston, D., & Kuehnle, K. (1981). *Procedures to develop and monitor progress on IEP goals.* Minneapolis: Institute for Research in Learning Disabilities, University of Minnesota.

Parker, R. I., Tindal, G., & Hasbrouck, J. (1991). Countable indices of writing quality: Their suitability for screening-eligibility decisions. *Exceptionality*, *2*, 1–17.

Salvia, J. A., & Ysseldyke, J. E. (2001). *Assessment in special and remedial education* (8th ed.). Boston: Houghton Mifflin.

Shapiro, E. S. (2004). *Academic skills problems: Direct assessment and intervention* (3rd ed.). New York: Guilford Press.

Shapiro, E. S., Keller, M. A., Lutz, J. G., Santoro, L. E., & Hintze, J. M. (2006). Curriculum-based measures and performance on state assessment and standardized tests: Reading and math performance in Pennsylvania. *Journal of Psychoeducational Assessment*, *24*, 19–35.

Shinn, M. R. (1981). *A comparison of psychometric and functional differences between students labeled learning disabled and low-achieving.* Unpublished doctoral dissertation, University of Minnesota, Minneapolis.

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

Shinn, M. R., & Shinn, M. M. (2002). *AIMSwebfi training workbook: Administration and scoring of reading maze for use in general outcome measurement.* Eden Prairie, MN: Edformation, Inc.

Skiba, R., Magnusson, D., Marston, D., & Erickson, K. (1986). *The assessment of mathematics performance in special education: Achievement tests, proficiency tests or formative evaluation?* Minneapolis, MN: Special Services, Minneapolis Public Schools.

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

Tindal, G., German, G., & Deno, S. L. (1983). *Descriptive research on the Pine County Norms: A compilation of findings* (Research Report No. 132). Minneapolis: Institute for Research in Learning Disabilities, University of Minnesota.

VanDerHeyden, A. M., Witt, J. C., & Naquin, G. (2003). Development and validation of a process for screening referrals to special education. *School Psychology Review*, *32*, 204–227.

Wayman, M. M., Wallace, T., Wiley, H. I., Ticha, R., & Espin, C. A. (2007). Literature synthesis on curriculum-based measurement in reading. *Journal of Special Education*, *41*, 85–120.

Witt, J. C., Daly, E. J., III, & Noell, G. H. (2000). *Functional assessments: A step-by step guide to solving academic and behavior problems.* Longmont, CO: Sopris West.