CHAPTER 1

What Is CBM and Why Should I Do It?

A third edition. Whew!

For those of you who are back from the first or second, we think you'll find that today's CBM is not what you remember it as. Excuse us a little if we get teary eyed for "the good old days" and reminisce about how CBM started and where it's come from. As CBM nears its 50th anniversary, there is a lot to examine and quite a rich and varied history to explore.

When we started this undertaking, we knew it was time for a new edition. Keeping up with the new publishers, new measures, and new applications was becoming increasingly complex. To include the changes that have been made since the last edition, we dug into researching all the advances, and the more we examined, the more we found that things have changed, and changed, and changed. So welcome! Hold onto your hat and enjoy the ride.

For those of you who are new to *The ABCs of CBM*, welcome to you too! This book is about an assessment tool called curriculum-based measurement (CBM). The book begins by explaining a little bit about what CBM is and where it came from because it is important to understand the context and application. But most of the book focuses on the nuts and bolts (i.e., the ABCs) of how to use CBM in a classroom, school, or district to improve the quality of educational decision making.

Given the number of assessment and evaluation initiatives present in education today, you might be wondering why you need to know about CBM. That is a legitimate question.

CBM was originally introduced as an *alternative* to other assessment procedures that teachers were doing—teacher-made tests and norm-referenced, often individually administered, tests (Deno, 1985). CBM was originally developed as a better way to monitor a student's progress from where they were to where you wanted them to be. Think of the change from using maps to using a GPS.

When using a map, it certainly helps to know where you are starting and where you want to end up. From there, you can plan a route in advance. But at any point as you are driving or running the route, you might not know if you were on the right path, how long you had gone, or how long you had left to go. You would stop and consult the map, try to determine where you were (which could be difficult if the street signs or landmarks were not up to date), and then plot a new route. If something changed the layout, like an accident, road construction, or a tree limb, you would be out of luck until you stopped and rerouted yourself.

Now that we have GPS, a whole new world is open to us! We can get real-time updates on traffic conditions, road closures, accidents, and even when to make sure to slow down because of police presence. That was what the invention of CBM was like. Before we might have had hand-drawn maps (i.e., teacher-made tests) or maps without the ability to react to current conditions. CBM was developed to monitor the route a student took from their current performance to where we wanted them to be (proficient at a certain skill).

During the 1990s, researchers began to discover that CBM could also be used to predict which students were on pace to master a skill and which weren't. As the emphasis on prevention and early intervention grew throughout the 2000s, CBM emerged as a key tool for universal screening too. Since that time, there have been several other advances in universal screening—especially with increases in computing power, computer adaptive tests (CATs) have really expanded too. In fact, you will see in the CBM suites we discuss that there is an overlap between the two. Some suites have CATs that are part of their broader family of measures (e.g., FastBridge and Star).

As the purposes that CBM serves have grown in popularity and importance (more on those later), states have begun to mandate this type of assessment. As of today in the United States, 40 states and the District of Columbia mandate universal academic screening assessment (Stollar & Dunn, 2025). This is most common in reading and in the earliest grades, K–3. In addition, 49 states and the District of Columbia require dyslexia screening as part of their early literacy initiatives (International Dyslexia Association, 2026). For most educators in the United States, academic screening, particularly in reading, is a legal mandate, with more states passing legislation each year. To learn if your state has passed legislation for screening, we recommend you visit your state's department of education website and search for *universal screening* or *screening for dyslexia*. Another resource is the COMPASS on The Reading League's website. For all the states that have already passed legislation, at least one CBM product has been approved for use. So the odds are, whether you've been teaching for decades or are new to the field, CBM will be a part of your professional life.

Why is CBM required for screening? For a variety of reasons, but four important ones are that CBM provides a reliable and valid way to identify:

1. if your core instruction is meeting the needs of most students,
2. which students need instructional support in addition to that core instruction,
3. which students are not making adequate progress given the instruction they are receiving (core plus supplemental or intensive), and
4. which students need addition diagnostic evaluation to pinpoint what they need instructionally.

WHY DO I NEED THIS BOOK?

This book includes a set of skills that lead to quality instructional decisions. It is about collecting and using information. Whenever we work at something important, it is best to develop a plan before we start and to check on our progress while we are working. This allows us to work in an intentional and thoughtful way. It defines what we are trying to accomplish, alerts us when what we are doing isn't getting us closer to that goal, and gives us the information we need to determine how to change. Educating children, adolescents, and young adults certainly fits within the definition of an important activity. Therefore, the process of education should include things like goal setting, planning, instruction, and monitoring. To do these things well, an educator needs information! The quality of the information you have will determine the quality of the work you do that is informed by that information (i.e., assessment data).

In the United States, for example, there are millions of school-age students with serious reading problems for various reasons—they may not have received high-quality instruction before, or they may have a disability. As a result, we have an increased responsibility to make informed decisions when working to teach important skills like reading and when tackling the needs of students who face problems learning. But it is difficult to think our way through these important efforts without something concrete to think about. CBM provides exactly the kind of functional information required to inform educational decision making. Therefore, this book is designed to teach you how to gather and use meaningful data with an emphasis on testing less and teaching more—conducting assessment only when the information is essential for making better instructional decisions.

WHAT IS CBM?

CBM is an assessment tool characterized by certain attributes. We explain these attributes shortly, but first you should know what CBM "looks" like. Like most things in education, what it looks like has evolved with the times and the use of computers. Traditionally, you would see a teacher and student sitting together for a brief period (e.g., 1–3 minutes) or an entire class working on sheets of math problems, taking a spelling test, or writing a story with paper and pencil. Now a CBM can look like that or like a student sitting at a computer, both student and teacher sitting at separate computers, or maybe a student and teacher sitting together for a longer period doing much more varied tasks. But some things remain the same.

A CBM is *usually* composed of:

1. standardized directions and guides for administration,
2. a timing device (ideally both counting up and counting down),
3. a set of materials, which might be printed pages, a flip-book, a computer or tablet, a verbal prompt, or even manipulatives,
4. standardized scoring rules,
5. standards for judging performance, and
6. record forms or charts, which could be print or web based.

The directions given are usually straightforward because they either ask the student to engage in a task that is not that different from something they would do during class (e.g., read a story, write a paragraph, or solve computation problems) or they are presented with clear practice items and explicit demonstrations. During CBM, as the student performs these tasks, they are usually timed in some way—either to calculate a rate of performance of a skill (i.e., how many correct in a minute) or to record how long it took to complete the task. Therefore, the person giving the test will typically have some sort of timer (unless the timer is built into the computerized assessment). Also, you will see the student's level of performance on a CBM charted on a graph, either by hand or more likely on a computer, so that trends in their learning can be analyzed over time.

When watching CBM, some of the tasks will look very much like a teaching activity (except without the corrective feedback). For example, oral passage reading (OPR) is the student reading connected text aloud. In CBM you see OPR in most classrooms, especially in the primary grades, because one of the supporting principles of CBM is an idea called *alignment*. The principle of alignment basically holds that your educational efforts will be more effective if you "test what you teach and teach what you test." *What* you teach is called the *curriculum*. It is the goals and objectives that must be met to achieve social and academic competence. (This is a fairly standard definition. The word *curriculum* comes from the Latin word *currer* for racing chariots. The *curriculum*, then, is the "course of study" to be followed on the way to the finish line.)

WHY WERE THE OTHER ATTRIBUTES, LIKE THE TIMING AND CHARTING, DEVELOPED?

CBM evolved out of work by Stan Deno and Phyllis Mirkin in the late 1970s and early 1980s at the Minnesota Institute for Research on Learning Disabilities (Deno & Mirkin, 1977). They were working on an intervention process called data-based program modification (DBPM). DBPM was a complete package of procedures for establishing goals, planning interventions (with a heavy emphasis on collaboration and consultation), and monitoring. However, in order for DBPM to work, there needed to be a continuous data collection system in place to produce the information needed to guide the decisions that fueled the program modifications. Continuous data collection was also needed because, as many of the instructional interventions were designed through consultation, the person delivering the lessons was not always the person responsible for the students' learning (just like today).

Deno and Mirkin realized that they needed an assessment system built on a set of common principles and composed of standardized procedures and rules. In a way, this sort of system already existed in the form of applied behavior analysis for areas such as classroom management and social behavior. But there wasn't a system like that for academic content. So Deno and Mirkin began developing CBM.

CBM is characterized by several attributes (Deno, 2003):

1. The first and most obvious is *alignment* (mentioned above). Within CBM the students are tested on the tasks and content they are being taught. This means sometimes the

content and structure of the teaching and testing tasks are exactly the same (such as reading a story) or they could include foundational skills or specific applications (think matching letters to their sounds). The stimulus materials the student is given for both types of tasks often look the same. They might include reading a printed story, writing a story on lined paper, or completing a page of math problems. But even if they aren't whole sets of materials, they are specific parts that represent important skills such as determining which of two quantities is larger or breaking words into their sounds. The responses the student is expected to make are also the same for teaching and testing, such as reading aloud, completing the math problems, or spelling dictated words. Or they could be specific skills that are important to performance of the task such as blending sounds into words, reciting math facts, or transcribing dictated words.

2. The measures are *technically adequate*, which means they must have established reliability and validity. For evidence of the reliability and validity of many curriculum-based measures, you can check out the National Center on Intensive Intervention (NCII) (*www.intensiveintervention.org*). Even though CBM is used within classrooms by teachers, it is *not* informal assessment! Informal assessments typically have not been shown to be technically adequate (that's one of the things that makes them informal, not a tendency for the assessor to wear jeans and a T-shirt).

CBM is an empirically supported process with substantial technical adequacy. Over the past 50 years, there have been hundreds of solid empirical research studies in excellent journals supporting the application of CBM. In fact, because CBM is used to summarize both a student's level of performance and their rate of progress, it has been examined in ways traditional measures have never been examined.

3. CBM typically makes use of *criterion-referenced standards* in addition to norm-referenced standards (we'll explain this later in this chapter).

4. *Standardized procedures* are used for administering and scoring a CBM. All those using CBM who want to share their data with others (e.g., as part of a program evaluation or a formal student report) must follow the same administration and scoring rules. For example:

- Standard tasks are used for each content area (e.g., a 1-minute timed oral reading of a passage at a certain grade level).
- Standard administration and scoring directions are employed for each procedure.

5. *Performance sampling* is used (producing what is sometimes called *behavioral data*). CBM procedures employ direct, *low-inference* measures through which correct and incorrect student behaviors, on clearly defined tasks, are counted within a set time interval (usually in minutes) although some tasks, particularly for younger students, are untimed. Therefore, inference and conjecture about the meaning of the resulting scores is kept to a minimum. For example, a reading CBM might tell you that the student read a fifth-grade-level passage at "47 words correct per minute with no errors."

6. *Decision rules* are put in place to provide those who use the data with information about what it means when students score at different levels of performance or illustrate dif-

ferent rates of progress on the measures over time. These rules are based on performance criteria and standardized through sampling or experimental procedures.

7. CBM emphasizes *repeated measurement* over time and can be used to identify rates of progress as well as levels of performance. Some measures are exclusively for universal screening and some are exclusively for progress monitoring, but many can be used for both applications. How often a CBM can be repeated depends on the measure. Some are designed to be given three times a year and can't be given more often than once a month, generally because they take a long time to administer. Others can be given once or twice a week—even daily!—though you don't often need that frequent a measurement. Other non-CBM tests, like state tests, can only be given once a year. A CBM is given more frequently to give teachers the information to make immediate adjustments to a student's educational program when needed. Because CBM also often measures what is being taught, and learning is a change in performance over time, these repeated measures illustrate the degree to which current instructional interventions are producing learning. As a result, the use of CBM for screening and progress monitoring allows educators to judge the quality of their own instruction and to decide when changes need to be made. Therefore, CBM data don't just help teachers decide *what* to teach, they can also help them decide *how* to teach. They are important tools for this type of evaluation.

8. CBM is also *efficient*. It is efficient in implementation because people can be trained to give the measures in a short period of time and most measures can be given quickly. When you use performance data, you draw conclusions directly from what the student did on a test. For most curriculum-based measures there is no need to convert the raw score, you can use that score to compare the student's level of performance and rate of progress. Some newer curriculum-based measures, however, use equating practices (we'll explain in Chapter 5) that don't use the raw score but have extensive behind-the-scenes transformations to put the scores onto a consistent scale.

An example from some of the initial work that is still in use today: If a student reads 47 words correctly per minute and the criterion for the passage is 60 correct words per minute, then the conclusion is that they are reading 13 correct words per minute less (i.e., slower) than they should. That's it! Many of the CBM measures are that simple to interpret (and even more importantly, that simple to convey to administrators, parents, and even our students).

For classroom purposes, CBM results are summarized and interpreted as simple performance statements and do not need to be converted into percentiles or normal-curve equivalents to be understood (although those other metrics are often available too). All you need to know is that the student reading 47 words correctly per minute must be taught to read 13 words correctly per minute faster than they are currently, or that 12 of your 30 students are currently reading below that 60 correct words per minute standard that you want them to meet by May. Faster is not the goal, but reaching the performance standard represented by a number is the goal because that "number" represents adequate reading performance. This is a very important point. There is a difference between *fluency* and doing something *fluently* (Hosp & Suchey, 2014).

9. Last, the CBM data can be *summarized efficiently* by using a variety of techniques ranging from pencil-and-paper charts to a web-based data management system. This efficiency makes the data immediately accessible at any level of the educational system—each student, subgroups of students, whole classrooms, entire grade levels, schools, districts, states, and more!

HOW IS CBM DIFFERENT FROM OTHER FORMS OF MEASUREMENT?

Many of the most important differences were spelled out in the nine attributes listed above. However, there are some fundamental CBM ideas that support those attributes.

Anyone who has spent time around education knows that there are all kinds of assessments available in schools. These range in structure from statewide accountability tests to simple handwriting rubrics. In education, we use these measures to inform our decision making. And the forms of these measures usually have to do with the functions they are designed to fulfill.

There are different forms of measures because there are different kinds of decisions to make and different ways to go about making decisions. CBM, as explained above, was designed to help teachers plan instruction and monitor outcomes to see if instruction is working. There are four ways the structures of CBM reflect this purpose: (1) by aligning with the *standards (or curriculum)*; (2) by measuring *alterable variables*; (3) by employing *low-inference measures*; and (4) by employing *criterion-referenced* decisions.

Standards/Curriculum

When we say a measure is "curriculum-based," we expect to see that measure sampling the tasks and content that students are taught. This might not be the case for measures based on ideas about general achievement, cognitive ability, or perceptual processing. Those tests may not be built to target the content a student is being taught. In fact, they may actually have been written to avoid it. This is also why these types of measures have little to no utility when it comes to planning instruction.

Curriculum is the term that was often used to describe *what* students were taught or expected to learn, with *instruction* being the *how* (that's why colleges of education have departments such as Curriculum and Instruction or Teaching and Learning). Today we more often use the term *standards*. Standards is a broader term that includes not just *what* the students are expected to learn, but also *how well*. What a student learns is important to know, but we also want to know how well we expect that student to perform that skill or to know that information. Although each state is required to develop its own academic standards across various content areas, there is substantial consistency nationwide. Initiatives like the Common Core State Standards, developed by the Council of Chief State School Officers and the National Governors Association, have significantly shaped this alignment. While some states adopted these standards in full, others used them as a foun-

dation to craft their own versions—ultimately influencing all 50 states and the District of Columbia.

Over time, standards have also grown more rigorous, reflecting advances in educational research and practice. For instance, decoding skills were traditionally taught in first grade, largely because kindergarten wasn't universally mandated. As studies have underscored the importance of introducing foundational reading skills earlier, many states have shifted decoding instruction into the kindergarten standards.

Since CBM was developed 50 years ago, it won't be changing its name to keep up with the times. Our guess is that you anticipated curriculum-based measures reflecting the curriculum (good for you!). But CBM is also designed to function within instructional systems that include systematic instructional interventions and student mastery of performance goals or standards (such as a multi-tiered system of supports [MTSS], also called response to intervention [RTI] frameworks). The evidence base of this kind of system indicates that direct measurement of student learning is most effective (Supovitz, 2012). Measures designed to function in other problem-solving paradigms, such as the traditional student-deficit model or those that assume that instruction should yield a normal distribution of skills, are designed differently. But how are they different?

Alterable Variables

One of the most important differences between CBM and other measures used in education and psychology is that CBM targets alterable variables. In education, an *alterable variable* is something that can be changed *through instruction.* Performance on curricular tasks is considered alterable because it is under the direct control of teachers (i.e., student performance can be changed through effective instruction). CBM was not designed simply to document the existence of problems or even to determine their cause. It was designed as a data collection system that would produce the information required to guide instruction. One of the things CBM can do very well, for example, is tell a teacher about the level of a student's knowledge about a particular skill. This information has immediate implications for instruction because instruction, by definition, is the provision of new knowledge.

This brings us to the issue of alterable and unalterable variables. There is considerable debate about whether measures of unalterable student-related variables (like home life, how previous siblings did, or parents' income level) provide useful information for guiding instruction. More to the point, students' skills *can* be increased by the teacher through instruction. However, things like a student's parents' income level or how their siblings did are not. Therefore, time spent measuring them, assuming the measures work, is time spent looking at things that teachers can't do anything about. Worse yet, even if measures of those variables work, the information they yield is still useless without good information about what skills a student needs to learn—so, in the end, CBM is always needed.

Low-Inference Measures

Tools that measure one thing so that conclusions can be drawn about something else require us to make *inferences*. Those that require us to process assessment results by way of some

theoretical applications are called *high-inference* measures. For example, a cognitive ability test (e.g., IQ test) does not have any cognitive ability items on it, but it does have items from which the test user is expected to make *inferences* about the student's cognitive ability. Therefore, while a student may assemble geometric shapes out of blocks on a cognitive ability test, the score is not reported in terms of "geometric shape production," but in terms of "cognitive ability"—the construct that the task infers. And we can't miss an opportunity to share why this is not helpful information. First, teaching a student to do puzzles will not increase their IQ, and second, IQ is a construct built on exposure and experience with middle-class culture, which is not very instructionally relevant or helpful.

The fact that CBM is designed to sample the observable student behaviors that occur in a classroom distinguishes it sharply from the high-inference measures that have little instructional utility. CBM was not developed to explain how learning does or doesn't occur. And it was not designed to conform to any particular theory about how students think, attend, remember, or process information. Therefore, inference and conjecture about what the resulting scores actually mean is kept to a minimum. A CBM employs direct (low-inference) observations during which correct and incorrect student responses to the tasks being taught (e.g., addition) are counted, often within a set time interval (e.g., in minutes). If the student works seven addition fact problems in 1 minute, their score is reported as "seven addition facts per minute." If the criterion for addition facts is 40 per minute, the seven-per-minute score is simple to interpret: It means this student needs instruction on addition!

Criterion-Referenced Decisions

Another way that CBM is different from most traditional educational and psychological measures is that it escapes the normative tradition and employs decisions that are criterion-referenced (although norms for many of the measures are also available). Criterion-referenced decisions are used to determine if students can demonstrate their knowledge of certain tasks at specified performance levels (i.e., criteria—one component of standards). The basic assumption is that students who do not know a skill and need instruction on it will do poorly on the measure of that skill. Whereas those who *do* know the skill will perform well.

One of the biggest problems with the utility of educational evaluation is that its history has been grounded almost exclusively in *normative standards*. There is nothing wrong with normative comparisons or the measures used to conduct them, as long as your goal is to find out how a student's level of performance compares to the performance of others. But that isn't the most important thing teachers need to know! For planning a lesson, it is more important to know if the student has or hasn't mastered the skills about to be covered (or what they need to be taught next). Knowing how a student compares to other students does not provide that information!

CBM came directly out of an intervention program and was designed to inform teachers' decisions about *what* and *how* to teach. As has already been explained, CBM was designed for instructional utility. This meant that the measures had to be:

- Aligned with standards/curriculum;
- Sensitive to instruction;

- Repeatable so that screening and/or progress monitoring could occur; and
- Criterion-referenced so that they could be used to determine when a student had mastered a task.

These conditions allow teachers to identify which students need additional instruction, set goals, determine the level of a student's prerequisite knowledge, align instruction with outcomes, and track progress toward goals.

WHAT ARE THE MAIN ADVANTAGES OF CBM?

If we have to pick a few advantages, we will go with efficiency, alignment, and repeatability. The first advantage, efficiency, is important because no one wants to use a measure that is awkward, confusing, or burdensome. Who has time for that? CBM is quite simple to use and understand. This means less time assessing and more time teaching. Although some measures are now untimed and can take up to 30 minutes to administer, there are an increasing number of computer-administered options, group administration options, and scoring that is either concurrent to administration or automated. A lot of these characteristics aren't specific to CBM, but coupled with the other advantages, they make CBM a good option.

The second advantage would have to be CBM's alignment, or linkage, with instructional outcomes. Alignment between measurement and the content being taught allows the user to make better decisions. For example, alignment improves decisions about what the student can and can't do. Most of the curriculum-based measures require the student to actually do what we expect them to in order to meet the standards (e.g., read from connected text, spell dictated words, or solve math computation problems). Teachers usually collect the same information during their instruction.

Alignment is often lost with traditional normative measures, as these are constructed by using a sample of items selected across a wide range of difficulty. This format starts with very easy items and moves quickly through increasingly complex material. Unfortunately, in order to cover a range of skills and keep such tests down to a manageable size, the *curricular distance* (i.e., where these skills fall on a continuum of complexity) between items on these tests is often large, and very few items are provided for each skill. Alignment is lost because of the limited number of items for each skill and because some skills must be completely left off the test.

Alignment is also lost when measures use item formats presenting the student with tasks different from those they actually need to use. For example, group-administered tests often ask students to identify answers by circling or matching them. In actual practice, students don't need to identify correct answers; they need to produce them! The two skills are different. Originally, all curriculum-based measures used production of skills. However, in recent years, there have been some advances that use multiple-choice formats. This is so that they can be computer administered and scored. While that affects the alignment of the measure to the standards, it increases the efficiency of its use, and as we are finding out, can be done without losing any of the predictive utility.

Our third advantage is CBM's repeatability. Typical normative achievement measures can't be used to decide if instruction is working within a fairly short period of time. They are designed to yield scores that are highly stable over time (a student's score on normative tests should not change across a semester or grading period), and they don't have enough alternate forms for frequent retesting. However, CBM allows for monitoring student progress either through intermittent screening (two to three times a year) or more frequently in a biweekly or weekly format. Frequent use coupled with alignment makes CBM more sensitive to instruction than typical achievement measures. This means CBM can be used to decide, within a fairly short period of time, when instruction is or isn't working. CBM can also be used to help one decide *how* to teach. It does this by letting us see, in a timely manner, if the instruction is working and/or when it should be changed.

By opening access to progress data, CBM supplies educators with a whole new range of information with which they can make an expanded set of informed decisions. Information collected during the process of instruction is called *formative evaluation*. Formative evaluation was a central component of the DBPM system originally developed by Deno and Mirkin. It involves the use of information from repeated direct measures to display trends in learning so that instructional decisions can be based on levels of student performance and rates of progress. Formative evaluation is, hands down, the most powerful tool available to any educator!

WHAT KINDS OF DECISIONS CAN I MAKE WITH CBM DATA?

As will be explained in Chapter 2, there are four major kinds of decisions we make in education:

1. *Universal screening decisions* to decide which students need additional instructional help and which don't;
2. *Progress monitoring decisions* to decide when to move on to new goals or modify instruction;
3. *Diagnostic decisions* to decide what kind of help a student needs; and
4. *Outcome decisions* to decide when special services can be discontinued and to document the overall effectiveness of efforts across all students.

The kinds of measures we use and the ways we use them depend on which kind of decision we are trying to make. As will be explained shortly, a *general outcome* and a *skills-based* CBM are often used as survey measures and a *mastery measure* CBM is often used as a specific measure.

HOW DOES CBM RELATE TO MTSS, RTI, OR DBI?

MTSS and *RTI* are terms that are often used interchangeably and often mean different things to different people. In general, they include the use of data-based decision making

for problem solving (data-based instruction, or *DBI*, is a term people often use to refer to this process). Key components of any good MTSS/RTI/DBI approach are the use of assessment for universal screening and progress monitoring decisions, as well as the provision of instruction and intervention in a tiered system such that individuals who have greater instructional needs are receiving more instruction and support. As mentioned above (and discussed in more depth in Chapter 2), CBM is an excellent way to make universal screening and progress monitoring decisions such as those central to MTSS/RTI/DBI. Because this book isn't specific to MTSS/RTI, we don't go into detail about other components such as the actual instruction or systemic application; however, there are several resources for you to check out about MTSS, RTI, and DBI in Appendix B.

HOW DOES CBM RELATE TO CURRICULUM-BASED EVALUATION?

The third key component of MTSS/RTI/DBI that we mentioned has to do with detailed decision making about student needs and learning. It is a hallmark of MTSS/RTI/DBI because students who are experiencing the greatest difficulties are the ones who need the most intensive teaching. Curriculum-based evaluation (CBE) is one approach to instructional decision making to meet the needs of students who are struggling. As you might have guessed from the whole "curriculum-based" thing, CBE is a systematic problem-solving process that relies heavily on CBM for the data on which we base our decisions (see Hosp, Hosp, Howell, & Alison, 2014, for the how-to of CBE).

SO, CBM DOES JUST ABOUT EVERYTHING?

Well, it doesn't teach!

CBM is *not* an instructional method or intervention. It is a tool for improving instruction that is compatible with diverse instructional approaches. Similarly, CBM is *not* a curriculum or set of standards. So, there isn't a CBM reading program.

CBM is a measurement overlay, which means the CBM administration and scoring rules are like templates that can be laid over goals and objectives from an assortment of content areas. This makes CBM uniquely valuable in situations where different teachers may be using different instructional methods or the same teacher may have different students being taught in different ways.

There are sets of published curriculum-based measures that have been developed around particular sequences of goals, but the tasks and goal sequences used in those measures are not the defining elements of CBM (they are defining elements of the different tasks, objectives, and curriculums/standards on which they are based). The defining elements of CBM are the curriculum-based procedures for designing, administering, and scoring measures as well as recording, summarizing, and interpreting the data that result from those measures. Therefore, you can't buy one CBM that will be useful for all subject areas

or in all classrooms; for that you would need a full suite of curriculum-based measures such as those provided by the various publishers referred to in this book.

ARE THERE DIFFERENT TYPES OF CBM?

A measure gets to be a CBM instrument if it is designed, administered, and scored according to established CBM procedures. Three types of CBM procedures have been described: general outcome measures, skills-based measures, and mastery measures. These all share the qualities listed above but may differ in design according to their purposes and the nature of the skills they are designed to test.

General Outcome Measures

General outcome measures (GOMs) are used to sample performance across several goals at the same time by using capstone tasks that are complex in the sense that they can only be accomplished by successfully applying a number of contributing skills. In this measurement format, the contributing skills (i.e., subskills) are not separated out for direct attention as they are in the skills-based measures and mastery measures we describe in the next section. Instead, success or improvement on the GOM is assumed to reflect the synthetic application of the contributing skills. In this sense, GOMs are holistic, while mastery measures in particular, are atomistic.

Probably the best example of a GOM is oral passage reading. For a student to be able to read proficiently, they must be able to use a variety of skills at the same time. Those include the skills required to use letters, letter combinations, blending, vocabulary, syntax, and content knowledge. As a student improves in any of these skills you can expect to see some improvement in their oral passage reading. As a result, using oral reading as the GOM relieves you of the need to monitor each of these subskills separately (whether they are taught in isolation or in combination).

There are several obvious advantages to GOMs. The first is that they dramatically cut down on the number of different measures one has to introduce, manage, administer, score, and track. Having four or five GOMs to cover the areas addressed throughout a year, a teacher can have their monitoring system for the whole year in place on the first day. The use of GOMs also recognizes the limitations of isolating subskills from the context in which they normally are expected to function. Any time you present tasks in a format that is different from the way they will usually be used (e.g., asking students to read nonsense words or the sounds of letters in isolation), there is the risk that you will lose validity. A final advantage is that visual displays of progress on a GOM will show longer acquisition slopes (i.e., skills that are taking longer to master), allowing adequate opportunities for progress monitoring and data-based instructional modifications.

For the reasons listed above, GOMs are especially useful for universal screening and progress monitoring to get an overview of level of performance. The primary disadvantage of GOMs is the downside of all general procedures: They are *general.* If your student's oral

reading is inadequate and you think you need specific information about their relative skill patterns, you will not get that information from a GOM. Another limitation of GOMs is that some content areas do not have a capstone task that represents the synthetic application of most of the content (especially one that is reasonably convenient to use). For example, GOMs are difficult to develop in mathematics beyond the early grades. This is why some math areas use skills-based measures.

Skills-Based Measures

Skills-based measures (SBMs) are designed to accomplish many of the functions of GOMs. They also have their particular advantages and disadvantages. Their main advantages are that they can be used to screen and progress monitor in content domains where capstone tasks are not available.

The best example of an SBM is probably math computation. At any particular grade level, math standards for computation are made up of a list of specific skills. For example, second-grade standards might include addition facts, double-digit addition without regrouping, double-digit addition with regrouping, and subtraction facts. There is no single task to demonstrate proficiency on all of these skills—each needs to be measured directly, using an SBM.

SBMs are constructed by first identifying the set of goals that will be taught within a curriculum area. The time frame you cover could sample goals for an entire year or for shorter periods. Once the goals have been identified, items are then prepared to assess each goal. The items for the same goal should be of equal difficulty. Next, the items are placed in random order (from the student's perspective) into a set of tests (in fact, they should be in a deliberate order, but one that is not readily identifiable). This produces a set of equivalent measures providing balanced coverage of the same content.

The items on these tests are not placed in the order in which they are taught or in order of complexity. All the items covering the same goal are not grouped together. Items should be arranged so that each goal is equally represented in each section (i.e., beginning, middle, and end) of the test. It is good to note which skill each item is measuring, however, so that you can link performance on the measure back to instructional objectives.

SBMs are generally administered by including directions like "Work as many items as you can. If you come to one you don't know, you can skip it." When given these directions and measures constructed as we have described them, students who are beginning to work on a set of skills will skip many problems and get lower scores. As they progress through the curriculum and learn new skills, their scores will improve because there will be more items they can work. Therefore, SBMs can sometimes be used to progress monitor, as they will produce long acquisition slopes like GOMs do. In addition, they can yield some analytical information as long as steps are taken to ensure that an adequate sample of each kind of item is provided and that the items are cross-referenced to goals.

One big disadvantage of SBMs is that when instruction begins, most of the items will be irrelevant to the student because they will be above their current level of performance. Near the end of instruction, most of the items will again be irrelevant because they will

have already learned them. Basically, this means that at any given time only a few items on the test will be directly related to what the student is currently learning. However, because it continues to sample items already learned it also measures retention, which is one big difference between MMs and SBMs.

Mastery Measures

The last type of CBM is the mastery measure (MM). MMs differ from GOMs and SBMs in several ways, mainly in the relative levels within the standards from which tasks are drawn and the relative sizes of the measurement net they spread. (The term *measurement net* refers to the size and nature of the sample a measure collects. For example, a test covering 25 computation skills would be casting a larger measurement net than one covering five skills.) GOMs present tasks that are relatively more complex and/or advanced than MMs; SBMs tend to cover more skills than MMs (i.e., they measure more by casting a wider net). Therefore, MMs are generally used on parts of the curriculum that contain discrete and easily identified sets (or domains) of items that are closely related by some common skill, theme, concept, or solution strategy. Examples of this sort of domain might include punctuation (for writing), multiplying fractions (for math), or sounds of letters (for early reading).

MMs are used in three situations:

1. When you really want to focus on a particular set of skills. These might include the so-called *tool skills*, which need to be performed at high levels of proficiency (e.g., letter formation, using the silent *e* to convert vowels, computation facts). Focus might also be important for skills that are pivotal to many other operations, like quickly going through the steps of multiplying fractions;
2. When you are trying to troubleshoot a problem and need to do specific-level testing (e.g., to see if a student is having trouble with reading comprehension because they don't know how to tell relevant from irrelevant information); and
3. To monitor learning when a skill is being taught in isolation. (It is important to note that, even if an MM focuses on an isolated skill, it does not mean that skill should be taught in isolation. The skill is measured in isolation only for purposes of focus.)

The disadvantages of MMs come with their narrow focus. They are not good for surveying general levels of performance or for monitoring growth on long-term goals. Using a series of MMs to progress monitor will produce a profile of closely packed peaks and valleys that look like the teeth on a saw blade (see Figure 1.1). This profile emerges because, as soon as a student starts getting high scores on one of the very specific measures, a new one is introduced, and their score goes back down. That is called a *measurement shift* (or as we sometimes like to call it, "jumping off cliffs"). A GOM or SBM covering what amounts to the same slice of curriculum covered by a series of MMs won't produce these measurement shifts and will provide the long classic learning curve needed for decision making (see Figure 1.2).

A summary of the attributes for each type of measure is provided in Table 1.1.

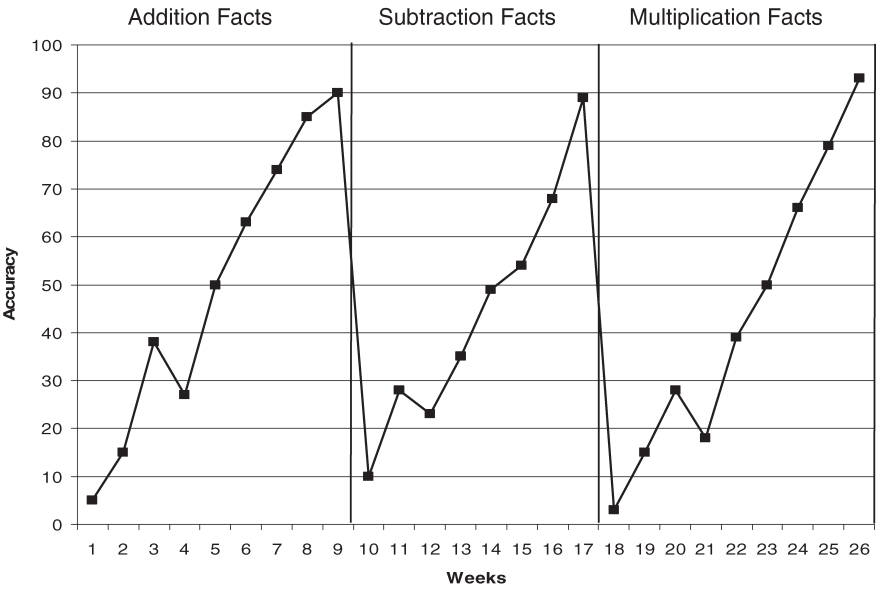

FIGURE 1.1. Example of an MM progress monitoring profile.

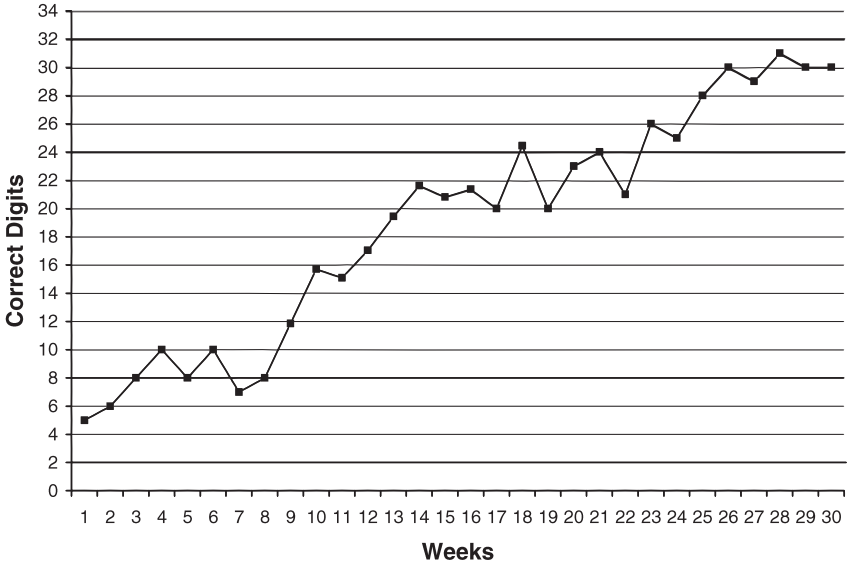

FIGURE 1.2. Example of a GOM or SBM progress monitoring profile.

TABLE 1.1. Comparison of the Three Types of Curriculum-Based Measures

General outcome measures (GOMs)	Skills-based measures (SBMs)	Mastery measures (MMs)
	Primary uses	
• Screening • Survey-level testing • Progress monitoring	• Screening • Survey-level testing • Progress monitoring	• Diagnostic evaluation • Specific-level testing • To target content areas of concern • To target different proficiency levels and response types
	Structure	
• Uses global/interactive tasks • Separate skills are not isolated or marked • Targets long-term goals • Often includes common classroom tasks	• Composed of mixed items drawn from a set of goals • Skills are usually sampled across a whole year's curriculum • Separate skills may be isolated or marked • Items are often cross-referenced to goals	• May only test one specific skill or short-term instructional objective • A large sample of performance is collected on each skill • Items are referenced to skills and/or proficiency levels • Some skills may be examined in isolation
	Advantages	
• Provides perspective • Gives an overall impression of skill level • Useful for monitoring • No measurement shifts • Illustrates retention and generalization	• Gives an overall impression of skill level • Useful for monitoring • No measurement shifts • Illustrates retention	• Useful for double-checking a problem indicated on a GOM or SBM • Useful for checking hypotheses about missing skills or subskills • Provides focus
	Disadvantages	
• Provides little diagnostic information • Doesn't provide information about specific skills • Often includes a high proportion of items that are either above or below the student's skill level • Some content areas don't have convenient capstone tasks	• Small sample for each goal limits diagnostic utility • Often includes a high proportion of items that are either above or below the student's skill level • May not require generalization or interactive use of the skill	• Doesn't provide the big picture (no generalization or application) • Skill–subskill relationships may not be real

Composites

CBM has traditionally used the scores from each measure individually, whether that's the raw score or a transformed score. However, in preschool and kindergarten researchers have found that there is no single GOM or even SBM that can be used as a single indicator for screening, and MMs are generally not robust enough measuring only a single subskill. To remedy these issues, many publishers use composite scores in the early grades and preschool. This approach provides a robust score that captures many skills in order to make more accurate decisions about which students need additional instructional assistance.

I HAVE NEVER SEEN CBM BEING USED— IF IT'S SO GREAT, WHY ISN'T IT MORE POPULAR?

In the first edition of this book, this was an important topic. Most educators likely hadn't seen CBM being used because it was just hitting the mainstream. Today, the more likely issue is something like "I've never heard of CBM, but I've been using DIBELS for my whole career." For the first 20-plus years of its existence, there was no CBM product (complete with briefcase and clever branding swag). As such, if you used CBM, you likely knew exactly what, why, and how you were using it. Today, there are several publishers that develop, provide, and/or sell products that fall under the umbrella of CBM. You're probably familiar with one or more of them. If you think you haven't seen CBM before, as you read through this book, we bet you will come across some specific publisher materials that will make you go "Oh! That's a CBM!" Insert V8 head slap here.

IS CBM USED WITH SPECIAL EDUCATION OR GENERAL EDUCATION?

CBM was originally used in special education because of its ability to target specific skills and its sensitivity to instruction, making it particularly useful for adjusting instruction to individual student needs. It was designed to monitor a student's progress over time to make more frequent instructional decisions. Eventually it was discovered that CBM was also really good at predicting which students were likely to be proficient in a skill at the end of the year and which weren't. However, special educators really aren't the only ones who do those things. The use of CBM by general educators has been growing over the past few decades. This is, in part, due to the expansion of MTSS/RTI as a service delivery model, but also because everyone has become increasingly aware of the need to conduct universal screening to catch those who are falling behind as soon as possible and to progress monitor to make instructional decisions in order to help those students catch up. It is particularly important to progress monitor in high-impact content like reading, oral language, written expression, math, and social skills (many of these areas are covered in this book). As a result, CBM is being increasingly adopted by whole school districts (and states) as a system for use with all students and by all educators.

WHO GIVES CBM MEASURES?

We will go into greater detail on this in Chapter 12, but the simple answer is it depends on why the measures are being given. Often curriculum-based measures are used three times over the course of the year to screen all students by looking at their level of performance in key skills like reading, math, and writing. An increasing number of these measures can be group administered. Group-administered measures take from 3 to 30 minutes, but this is for a whole class or for each student working independently at a computer. Many measures, especially for younger students, require individual administration. If the flow of students and materials is managed smoothly, one person should be able to administer the necessary measures to an individual student in 5 to 20 minutes.

Giving a CBM for the purpose of analyzing learning problems is a different matter. This book is not about diagnostic assessment. Assessment for diagnostic purposes is usually carried out by someone who is an expert in the content area of concern as well as CBE, DBI, or a similar data-based decision-making system. This person could be a general education teacher, special education teacher, interventionist, content-area specialist, or school psychologist. When a CBM is given for analysis of a learning problem, specific measures are selected to check on the presence or absence of those skills suspected to be causing the problem. To do this, we need to have a set of these measures available. Publishers are creating an increasing number of CBM tasks that are more diagnostic in nature (using MMs), especially in early reading, early numeracy, and math. Although we focus on universal screening and progress monitoring, the publisher whose suite of materials you use might have others designed for diagnostic decision making.

SO, IF I WANT TO USE CBM, I DON'T NEED TO MAKE TESTS OUT OF THE INSTRUCTIONAL MATERIALS I'M USING?

No. Because one of the hallmarks of CBM is technical adequacy, it is best to use published or otherwise openly available materials that have been evaluated. For universal screening and progress monitoring decisions, reliable and valid instruments are a must. Enjoy having someone else do that work.

You don't need to have tests using the same formats and examples as the instructional materials. This is in line with why we use the term *standards* more than *curriculum*. You want tests that address the same skills (the content of the standards) and predict important outcome measures (a level of proficiency for the standard) without using the exact same materials. With so many publishers offering CBM suites these days, you have a variety to choose from. Standardization of those materials so that they are applicable across a variety of instructional programs, sets of standards, or sequences of topics, is key to their utility for three main reasons:

1. *Instructional programs don't follow the same sequences and schedules.* One of the biggest challenges in education is that what is taught and when it is taught really is not standardized across schools (even within a state where all schools are addressing the same

standards at each grade level). It is often not consistent across states, and it certainly is not standardized across published instructional programs. Obviously, this creates major problems in a society wherein students are mobile across schools, districts, and states.

2. *Program-specific tests may not tell you if the learning has generalized.* If you measure skills without being bound to any particular set of instructional materials, you are more likely to have better generalizability of your decisions. Many states are now providing guidance on a short list of instructional materials that have been vetted to meet the state standards. In fact, some may actually prefer to use CBM items that are somewhat different or at least mixed to try to ensure that learning has generalized (you don't want a student who can only work problems that are presented in a certain format).

3. *Program-specific tests will make the teacher dependent on the program.* Instructional materials do not remain constant. They are often revised, or teachers and schools select new ones. If teachers use program-specific tests, they will have to produce new ones every time a new program is selected. Wouldn't you rather put your energy into teaching?

WHERE DO I GET CBM MATERIALS?

There are many sources of CBM materials, and we present many of the ones on the market today. Most must be purchased, and others are free. Materials for the content areas addressed in this book are referenced in those chapters. Just remember that when you select a publisher's materials for CBM, you should be deliberate because it is an investment in time and money and you are trying to do the best for your students.

The NCII has a set of tools charts for academic screening, academic progress, behavior screening, and behavior progress (see *www.intensiveintervention.org*). These charts show evaluations of various assessments that have been conducted by groups of experts that serve on their technical review committees. We have been members of these committees from their beginnings (2003 for academic progress, Michelle; and 2008 for academic screening, John) and are dedicated to the crucial information that they provide. As you explore their site, you will find other materials that can help you expand your practice in other ways too.

Also remember that there is more to evaluation than giving a test. You have to score it properly, record the data accurately, and interpret them correctly. This book gives you information about CBM scoring rules and the interpretation of scores. Assessment is carried out to inform decision making; we need to know what the scores mean to use them.

WHERE DO WE GO FROM HERE?

Our goal for this first chapter of the book has been to answer some of the fundamental questions that are asked about CBM in general terms of the "whats" and "whys." Chapter 2 provides additional detail, putting CBM into a broader framework of decision making in education. For the rest of the book, we turn to the "hows," as in "How do I implement CBM?"

Chapters 3 through 10 each provide, for different content areas, a rationale for using CBM, a list of materials needed and where to get those materials, directions, and scoring procedures. Additional information that used to be in these chapters (i.e., how often it should be administered, how much time it will take to administer and score, information about the different types of CBM scores, how to write IEP goals and objectives, and frequently asked questions) is now included once (either in Chapter 2 or in Appendix A), since it was largely redundant. Chapter 11 takes you through the process and procedures for setting goals and graphing the data, and it describes how CBM fits into an MTSS/RTI model. Chapter 12 provides a guide for how to use CBM, how to get it going, and how to sustain it.

Appendix A provides answers to many frequently asked questions (FAQs). Appendix B provides a list of references cited in the chapters, resources that we find useful for either implementation or background information, and further readings to extend your knowledge on a topic. They're all merged into a single alphabetical list so you don't have to know which one we think it is.

The chapters are structured this way to serve as a reference when you are implementing CBM. When using CBM materials, it is important to know about the administration directions, scoring, and standards that are specific to those measures. Graphing and setting or writing goals is a similar process no matter what the content area, so it's like using CBM as a measurement overlay; those processes are provided in a way that can be applied to any content area.

So hold on to your hats and get ready to explore the ever-changing world of CBM!