CHAPTER 1

Introduction to School-Based Behavioral Assessment

School personnel engage in two core activities to support student success: **prevention** and **intervention**. By delivering high-quality core services across academic, social, emotional, behavioral, and physical domains, school personnel help students build the skills necessary to succeed and successfully overcome obstacles to success, thereby acting to prevent problems. By being responsive to the needs of children who require additional support in order to succeed through intervention, school personnel attend to the rights of children to a free and appropriate public education in the least restrictive environment possible. Although the foundational tasks of prevention and intervention are likely self-evident to the adults charged with successfully supporting children's well-being, the frameworks that have been used to enact those practices have varied over time.

Prevention and intervention are core activities for supporting student success.

Historically, intervention services in schools have been provided to students through a "refer-test-place" model, with a focus on individual students and their patterns of skills and deficits. For instance, a teacher might identify a student who regularly yells at their peers and refuses to return to class during passing periods, notice a concurrent drop in their grades, and suggest that this student be referred for an alternative class placement or even special education services. After the school support team and the student's parents agree to evaluate the student, the student's data would be collected and discussed to determine whether or not they meet the criteria for a specific designation and/or placement. If the student met the criteria for a behavior disorder, for instance, interventions could be mandated for them within an individualized education program (IEP) and might include a combination of strategies, such as individual time with the school counselor, small-group social skills training, and a behavior contract. The student's goals would be reviewed annually by the team, with the expectation that these interventions will be effective in helping the student to better access the curriculum and eventually succeed in school and beyond.

Viewed in isolation, this process for determining how to intervene with students appears as if it could be effective. If a student exhibits a problem, stakeholders collaborate to determine the specific nature of the problem and brainstorm ways to address it. Unfortunately, this process neglects any consideration of how best to support students; all students are provided with a general education curriculum, and individuals who do not respond to that curriculum are evaluated to determine if they meet the criteria for a disability and are, therefore, eligible for special education services. If the student is not deemed eligible, then they might receive additional support, or they might not. Important concerns, such as how other students are behaving in the hallway, whether students receive high-quality core behavioral instruction that appropriately addresses more challenging situations, or whether students are reinforced for behaving as expected, may or may not be discussed as well during these conversations. When the focus of the intervention and the data collected to evaluate the need for intervention rest entirely upon an individual student or situation, school personnel may miss critical information that informs supports for all students and may bear directly on the environmental and contextual factors influencing that individual student. *A wider view of possible targets of assessment may allow us to engage in proactive or preventive efforts that serve more than a single individual—and even decrease the amount of individual assessment and intervention that we have to do.*

WHY FOCUS ON SCHOOL-BASED BEHAVIORAL ASSESSMENT?

This book is about school-based behavioral assessment. *Behavior* involves social (e.g., relating with others), emotional (e.g., affections, feelings), and behavioral (e.g., observable actions, conduct) domains. S*chool-based behavioral assessment* refers to those assessment practices that can be applied within school settings to provide information about student functioning in social, emotional, and behavioral domains, with the intention of supporting student success and well-being.

The reasons why behavior is a core issue for prevention and intervention in schools have been made clear throughout the research literature: *Behavior matters for students, for the adults they become, for their parents, for their teachers, for their peers, and for society and the world at large.* Teachers report that classroom management and teaching students with special needs are among their top concerns for additional training (Wei, Darling-Hammond, Andree, Richardson, & Orphanos, 2009), but only 27% of university special education programs include a course on classroom management (Oliver & Reschly, 2010), and only 22% of elementary school teachers report having received adequate classroom management training (Wagner et al., 2006).

School-based behavioral assessment provides information about student *social, emotional,* and *behavioral* functioning.

Students with behavior problems are more likely than students without behavior problems to experience academic failure (Masten et al., 2005), exhibit "internalizing" problems like depression as adults (Kosterman et al., 2010), encounter the criminal justice system at an earlier age (Patterson, Forgatch, Yoerger, & Stoolmiller, 1998), and be unemployed after high school (Wagner, Newman, Cameto, Garza, & Levine, 2005). Indeed, the outcomes for students with emotional disturbance, a disability category that includes students with sig-

nificant behavior problems, are very negative for both the short and the long term. Half of all students identified with emotional disturbance drop out of high school, compared with 30% of all students with disabilities (U.S. Department of Health and Human Services, New Freedom Commission on Mental Health, 2003), and less than one-third are employed after leaving school (Wagner et al., 2005).

Evidence-based interventions can prevent and/or remediate the significant and long-term negative outcomes observed for students exhibiting social, emotional, and behavioral challenges. With the knowledge that students *can* be taught to demonstrate positive behavior and that these changes *can* benefit students, educators are faced with the question of *how* best to support positive student behavior in schools. For decades, many of us have thought that effective strategies available in the behavior management toolbox consisted of sending students out of class, giving them an office discipline referral, and possibly suspending or expelling them. This viewpoint is often accompanied by the idea that appropriate behavior should be expected and should not have to be taught in schools—thus, the school's job is to punish a child into behaving appropriately.

Although viewpoints have shifted toward more positive and proactive approaches, punitive and exclusionary practices persist today. More than 106,000 public school students received corporal punishment (physical punishment intended to inflict pain) in 2013–2014 across the 19 states in which such practices are still legal (Office of Civil Rights, 2017). A nationwide survey of state-funded prekindergarten programs suggests that 6.67 of every 1,000 preschoolers were expelled during the 2003 and 2004 academic years (Gilliam, 2005); a significantly higher number of expelled preschoolers were boys and were Black. Critically, although this survey highlighted a number of disturbing trends in our nation's preschools, it also identified one potential factor for mitigating these expulsion rates: preschool teachers who had access to behavioral supports from a mental health professional were significantly less likely to expel a student than those without such support. Although this finding is correlational and not causal, it supports a growing movement to integrate social, emotional, and behavioral support services within a comprehensive student health framework in our nation's education system.

Given the significant negative outcomes associated with unremediated social, emotional, and behavioral problems, and the societal and very human implications of a primarily punitive behavior management approach, we choose to interpret behavior through a different lens. **We believe that behavior is learned, understandable, and serves a purpose for the student**. If Lisa bangs on her desk repeatedly over the course of a school year, and when she does so, her peers look at her angrily and her teacher scolds her, this is a behavior that Lisa has learned: when I bang on my desk, people pay attention to me. Critically, we believe that Lisa was taught this behavior by her teachers and her peers; after all, they are the ones providing the attention! Desk banging is a behavior that works for Lisa; it doesn't take much effort, it is easier than using a full sentence that might be difficult to articulate or than raising a hand that might not be called on, and it's effective because it's really hard to ignore someone banging on their desk.

Behavior is learned and understandable and serves a purpose for the student.

In order to determine why Lisa may be exhibiting this behavior, the extent to which it is typical for other students her age, how often it is currently happening, and whether an intervention we have tried has made a meaningful difference in changing her behavior, we

need information that is geared toward answering each of these questions. In other words, we need to collect data. Unfortunately, in an educational landscape in which schools are charged with doing more for their students within traditional academic and expanding social, emotional, and behavioral domains, it is critical to recognize that for some school personnel, the word "data" has joined the more traditional "four-letter words" in the English language. When approached without respect for an educator's time and available resources to follow through with a plan of action, data collection can be just "one more chore" added to educators' already overflowing plates. Thus, at the core of this book lies our premise that *schools should be proactive in identifying and responding to student needs with reliable and valid data in order to make defensible decisions. However, this focus on prevention must be balanced with the critical need for efficiency and ease of use of data collection.*

School-based behavioral assessment practices must balance defensibility with usability.

WHY ENGAGE IN DATA-BASED DECISION MAKING?

As noted throughout each chapter, the process of data-based decision making informs this book. As described in detail in Chapter 2, we use the Burns and Gibbons (2012) model of data-based decision making, which itself was adapted from a model proposed by Bransford and Stein (1984). We briefly introduce the idea of data-based decision making by asking two major questions when considering how data are going to be collected: *What question am I trying to answer?* and *Which data will best answer that question?* These questions are answered with an additional question about available resources, that is, *What resources are available to collect data?* Although the first two questions may appear simple, it is absolutely critical that we answer them *before* we collect data. All too often, practitioners and researchers alike start with a method and then work backward to determine what questions the data from that method are answering. Many, if not most, school-based practitioners will identify with a scenario in which they are presented with a binder full of data and asked, "What are we going to do with all of this information, and what does it mean?" We too have been part of numerous school- and district-level data team meetings in which we are presented with spreadsheets full of student and school outcomes and are asked to figure out what are the next steps to take. If our assessment process starts with the instrument itself rather than with the questions we are trying to answer, we will find ourselves staring at a report and trying to work backward to make sense of it all. If we instead start by asking *What information do I actually need in order to solve this problem?*, we will collect the data that are relevant to our question and potentially save time and resources, and avoid needless effort.

Data-based decision making involves asking two major questions: first ask the reason for assessment, and then ask which data are needed given the available resources.

The type of data that we collect is not just driven by the method used to collect them, but also by the level of inference we are planning on making with the information. For instance, if the intent of an assessment is to identify students who are at risk with regard to a particular behavior (e.g., disruptions in class) or to provide an evaluative statement about

the effects of an after-school program to enhance prosocial behavior, the desired level of assessment generalization (or in some cases specificity) should be considered (Riley-Tillman, Kalberer, & Chafouleas, 2005). Is the purpose of the assessment to provide an aggregated statement about a child's behavior (generalization), or to gather information about a particular behavior at a particular time or in a specific setting (specificity or directness)? The reason for needing the data not only drives the type of assessment that is called for, but also the methods and practices used in the assessment.

What Question Am I Trying to Answer?

Although assessment tool selection should be guided by a multimethod, multisource, and multisetting perspective, making decisions is not a matter of simply adhering to this guiding principle— particularly as related to social, emotional, and behavioral domains. For behavior in particular, context matters in the expression of (what the student is doing) and the perception regarding (how it is being interpreted) what is being exhibited (Dirks, De Los Reyes, Briggs-Gowan, Cella, & Wakschlag, 2012). Thus, it is often appropriate to use multiple assessment methods (e.g., direct observation or rating scales) in collecting data from multiple sources (e.g., teachers, parents, and peers) across a variety of settings (e.g., cafeteria, hallways, and restrooms). One limitation of this approach, however, is that more does not always mean easier or better. Collecting lots of information can be time consuming and cumbersome, and the quality of the data might therefore be affected. Furthermore, in our field, we have not resolved how to reconcile conflicting data from distinct sources (Dirks et al., 2012). Thus, a balance must be achieved between collecting enough information to understand a problem situation and develop an intervention plan and ensuring high quality, accuracy, and relevance. For example, a teacher's daily homework record that is maintained in the classroom grade book may be more efficient and just as relevant as a daily written log that is completed by the student, parent, and teacher that duplicates the recording of homework completion. In this example, some precision might be lost with regard to understanding the context in which homework is or is not being completed; however, feasibility may outweigh precision. We might also consider a multimethod approach, using both sources of information to gain information on work completion; we could compare the results and consider the implications of the data from both sources that agree or how to move forward if the data disagree. In summary, we highly recommend that multimethod, multisetting, and multisource assessment practices be given priority; however, each assessment situation should be evaluated carefully to maintain precision in the context of quantity that could diminish usability.

The chosen assessment tools should provide relevant information about the target behavior(s), the contexts in which the behaviors are observed, and the distal events that, although not in the immediate context, may affect the occurrence of the target behaviors in the problem context. For example, attendance data likely will not help inform intervention decisions related to increasing proactive skills on the playground. Riley-Tillman and colleagues (2005) referred to the match between the assessment tool and the behavior and context as "goodness of fit." Not all assessment tools appropriately measure the same behaviors. As an example, behavior rating scales that assess a student's general state or status relative to a wide range of behaviors likely would not be useful when rating "out of seat" behavior

events per hour during math instruction. Similarly, within a class of tools such as Systematic Direct Observation, a good fit between what behavior is measured and how to measure it must be considered. For example, a direct observation duration recording method (e.g., What percent of class time did the student spend out of seat?) would yield different data than an event recording method (e.g., How many times during a class period was the student out of seat?).

Which Data Will Best Answer the Question?

Determining the type of decision to be made guides selection of the assessment tool(s). High-stakes decisions, like curriculum adoption or individualized behavior support planning, require accurate and relevant data that provide users with a high degree of confidence. Confidence is related to the degree of inference needed when interpreting the data. One indicator for determining directness is the extent to which the collected information is removed in time and place from the actual occurrence of the behavior (see Cone, 1978). For example, comprehensive behavior rating scales are considered indirect because the information is collected by another person, who responds to items based on a retrospective perception of the student's behavior. In contrast, Systematic Direct Observation is considered direct, in that the assessments occur as the behaviors are observed. As a general rule of thumb, "high-stakes cases" (e.g., a serious disruption or potential harm to the student or others or a consideration of change in placement to a more restrictive setting) should include a combination of direct assessment tools. In general, the more direct the measure, the more resource intensive the data collection can be—but this is a good example of a reasonable rationale for allocation of intensive assessment resources.

What Resources Are Available to Collect These Data?

As previously stated, although asking the first two questions serves an important function in identifying the problem to be addressed and the potential data that can facilitate a solution, the available resources are also an important consideration. Determining the resources required to collect and interpret data is equally necessary when selecting assessment tools (i.e., how feasible is it to collect the data in a given situation?). Feasibility refers to a consideration, for example, of the time needed to train someone to accurately use the tool, of the intrusiveness of using the tool in the required setting, of scheduling a time for data collection, of the complexity of using the tool, and so forth. For example, asking a teacher to monitor student behavior every day for a full semester may not be possible given the class size, the training required, and fluency with assessment, given other instructional responsibilities. Together, these examples help to define the overall usability of the proposed assessment plan. For example, the factors that influence usability of an assessment could also include acceptability, family–school collaboration, system support, system climate, and understanding of its instrumentation and procedures (see Briesch, Chafouleas, Neugebauer, & Riley-Tillman, 2013). In summary, it is important to consider that resources are finite, and that wasting them on assessments that go beyond what is needed to answer a question can be a mistake that is just as problematic as not providing adequate resources for assessment in the first place.

PURPOSE OF THIS BOOK

Behavior matters: for students, for the adults that they become, for teachers, for parents, for schools, and for communities. Schools need efficient and effective ways to ask questions, approach problems, and determine solutions that support student success in social, emotional, and behavioral domains. Throughout this introductory chapter, we have provided initial considerations for school behavioral assessment guided by the conviction that (1) schools can and should work to support positive student behavior, (2) data are necessary in order to make effective decisions about what is or is not working, and (3) assessment practices must balance defensibility and usability. In Chapter 2, we introduce multi-tiered systems of support (MTSS) as a critical framework for integrating data-based decision making into supports that can effectively meet the needs of all students by expanding specific assessment considerations. Consistent with MTSS frameworks, our focus throughout the entire book is on school-based assessment as aligned with screening and progress monitoring purposes of assessment. As such, the reader should consult other sources for detailed information related to the functional, diagnostic, or evaluative purposes of assessment. Chapter 2 also describes how to drill down to and conceptualize the target behavior of interest that you'll be assessing. Chapter 3 examines the what, why, and how of extant data, which are data created as a byproduct of some other ongoing activity that can be used to assess student behavior. Chapter 4 describes the wide-ranging options utilized in Systematic Direct Observation, which has historically been referred to as the "gold standard" for directly observing student behavior yet might not always be the best choice in terms of usability. In Chapter 5, we introduce Direct Behavior Rating as an assessment option that relies on observers' perceptions of behavior during a larger observation period. Chapter 6 describes traditional behavior rating scales, wherein individuals provide retrospective ratings of student behavior across multiple items. Chapter 7 focuses on data integration, which involves taking the information derived from multimethod, multitrait, and multisource assessments and bringing it all together. Finally, in Chapter 8, we introduce a framework to guide you through conducting behavioral assessment in your own context and case studies that demonstrate practical applications.

For students or school personnel who are new to school-based behavioral assessment, we recommend reading the book all the way through in order to gain a more complete understanding of the various pathways decision making in assessment takes. For those readers who are approaching the book with some background knowledge in behavioral assessment and perhaps have an interest in specific behavioral assessment methods, we would still emphasize reviewing Chapters 2 and 7, given that assessment decisions that are defensible and usable undergird good decision making. By the time you have reached the end of this book, readers who engage in school-based behavioral assessment should have the knowledge needed to guide defensible and usable assessment decisions and an understanding of the tools they can use to make those decisions about student social, emotional, and behavioral needs.