CHAPTER 1

The Role of Direct Observation in School-Based Settings

In a typical school day, we are likely to see, look at, watch, glimpse, or view hundreds—if not thousands—of different student behaviors. We think of school-based observation as a decidedly different activity; however, what is it that distinguishes observation from other forms of seeing the world? In Sir Conan Doyle's (1892) short story "A Scandal in Bohemia," Sherlock Holmes explains to Watson the difference between seeing and observing while settling into his armchair:

> "You see, but you do not observe. The distinction is clear. For example, you have frequently seen the steps which lead from the hall to this room."
>> "Frequently."
>> "How often?"
>> "Well, some hundreds of times."
>> "Then how many are there?"
>> "How many? I don't know."
> "Quite so! You have not observed. And yet you have seen. That is just my point. Now, I know that there are seventeen steps, because I have both seen and observed."

The point that Holmes was trying to make was that although seeing can be done automatically and without conscious attention, observation is purpose-driven. In fact, the definition of observe is to "watch carefully, especially with attention to details or behavior for the purpose of arriving at a judgment" (*Merriam Webster*, n.d.). This definition implies that although we can *watch* someone or something passively and without an expressed intent, *observation* is seen as a means to an end—as helping to answer a specific question that has been posed.

Fortunately for us, there is not simply one way of observing (if there were, we could certainly end this book very quickly!). Herbert (1970) suggested that there are two types of observers: the realistic artist and the scientist. Both enter a setting with the goal of describing what is occurring; however, the approaches that they use to create these descriptions vary decidedly. The artist is concerned with creating a vivid, realistic picture of nature without concern for numerical judgments. Flaubert wrote, "When you pass a grocer sitting in his doorway, a porter smoking his pipe, or a cab stand, show me that grocer, and that porter, their attitude and their whole physical aspect . . . their whole moral nature, in such a way that I could never mistake them for any other grocer or porter" (in de Maupassant, 1902, p. lxi). The scientist, however, is concerned with quantifying what he or she observes and therefore must employ procedures that are reproducible. As Selltiz, Jahoda, Deutsch, and Cook (1959) wrote, "Observation becomes a scientific technique to the extent that it (a) serves a formulated research purpose, (b) is planned systematically, (c) is recorded systematically and relates to more general propositions rather than being presented as a set of interesting curiosa, and (d) is subject to checks and controls on validity and reliability" (p. 200).

Neither the approach of the artist nor the scientist is universally correct or incorrect; each simply provides the observer with information to answer a different set of questions. Depending on the situation, we may therefore be more apt to take the approach of the artist, the scientist, or, more likely, a blend of the two. When a friend learns of an upcoming dinner at one of the newest restaurants in town and asks us to provide a "full report," we are more apt to take the stance of the artist. Purposefully directing our attention to the décor, the clamor, and friendliness of the wait staff enables us to be able to paint a vivid picture of the restaurant experience. On the other hand, when we are approached by a classroom teacher with concerns regarding a student's behavior and asked to assist in identifying appropriate supports, taking the stance of the artist would be inappropriate. Within school-based assessment, wherein the goal is to answer specific, quantifiable questions about behavior (e.g., how often it occurs, the degree to which it disrupts the environment), a more precise, scientific approach is necessary in order to generate information that is objective and trustworthy. This book is intended to serve as a guide to conducting quality scientific observations that will produce maximally useful information regarding student behavior.

HISTORY OF DIRECT OBSERVATION IN SCHOOL-BASED ASSESSMENT

School-based observation is by no means a new concept. In fact, examples of observers entering classrooms to obtain objective estimates of student behavior can be traced back to the early 20th century. In 1914, Horn devised a system for measuring the participation of individual students in class discussion. Each time that the teacher asked a question, circles were marked on the classroom seating chart and squares were used to indicate a student response. Although other examples of observation systems can be found in the educational literature published over the course of the next few decades (e.g., Medley & Mitzel, 1958), it was not until the middle of the 20th century that direct observation came to be seen as an

important method for gathering information in school settings. The primary reason for this shift was the demand for better ways of assessing teacher effectiveness. In the early 20th century, literally hundreds of studies had been conducted in order to identify which characteristics differentiate effective from ineffective teachers. Typically, students were asked to identify their best teachers and then asked to describe the characteristics of these teachers (e.g., patient, polite, attractive) or demographic data were collected from the teachers themselves (e.g., intelligence, age). Unfortunately, when Domas and Tiedeman (1950) eventually undertook a massive review of more than 1,000 such studies, they found no relationship between teacher characteristics and student achievement variables. Given that determining what makes an effective teacher was still of great interest, a different approach to answering the question was clearly needed.

Beginning in the 1950s, educational researchers turned to the use of direct observation as a way of gathering information not just about *who* teachers were but also about *what they were doing* in classrooms. As Medley and Mitzel (1963) wrote, "Certainly there is no more obvious approach to research on teaching than direct observation of the behavior of teachers while they teach and pupils while they learn" (p. 247). Small armies of trained observers, equipped with notebooks and pencils, were sent into American classrooms to record the behaviors of both teachers and their students. One of the more well-known investigations of this era was led by Kounin (1970), whose team conducted observations in nearly 50 first- and second-grade classrooms. The results of this large-scale case study suggested strong correlations between desirable student behavior (i.e., work involvement, freedom from deviancy) and a number of teacher variables, including *with-itness* (the awareness of the goings on of a classroom), smoothness (the ability to move fluidly from one lesson to the next), momentum (the ability to maintain the flow of a lesson), and group alerting (the use of cues and signals to maintain student focus). Throughout the 1950s and 1960s, researchers designed observation coding systems to measure everything from the type of instructional materials teachers used (Medley & Mitzel, 1958) to the nature of verbal interactions (e.g., Wright & Proctor, 1968) for use within both research and applied contexts (e.g., teacher evaluation). Throughout this time period, however, the focus of classroom observation remained primarily on describing teacher behaviors.

In the second half of the 20th century, a change was occurring in the larger world of psychological assessment that would pave the way for the broader use of direct observation. For much of the early 20th century, psychological assessment emphasized a personological approach (Hayes, Nelson, & Jarrett, 1986). That is, explanations for human behavior were sought inside—rather than outside—of individuals. It was believed that psychopathology resulted from personality traits or characteristics that were relatively stable across different situations, and that understanding these traits was essential to being able to predict future behavior. Projective techniques (e.g., drawing, apperception tests) and personality inventories (e.g., Minnesota Multiphasic Personality Inventory) were therefore the mainstay of traditional psychological assessment and were used as a means of describing what a person *has* (Goldfried & Kent, 1972). The problem, however, was that an increasing number of studies began to suggest that behavior was largely context dependent. This research dated back to the 1920s, when Hartshorne and May (1928) placed children in a number of dif-

ferent situations in which they had the opportunity to be dishonest without being detected (e.g., opportunity to copy from an answer key on a test, cheating on an athletic task). What they found was that children were not universally honest or dishonest; rather, their behavior varied depending on the context. The need to take environmental variables under consideration in explaining human behavior became increasingly obvious.

Rather than believing that explanations for human behavior resided exclusively within the individual, behavior began to be seen as also shaped by situational variables. With the rise in popularity of applied behavior modification procedures in the 1960s, including the founding of the *Journal of Applied Behavior Analysis,* came an increasing interest in assessment procedures that would take these situational variables into account. The focus of assessment therefore shifted from understanding what a person *has* to understanding what a person *does* in different situations (Goldfried & Kent, 1972). Although several different methods can be employed within a behavioral assessment approach (e.g., self-report, analogue role plays), direct observation was to play a central role in understanding how individuals interact with the environments around them.

National surveys conducted of school psychologists help to illustrate how this shift influenced the nature of school-based assessment practices over time. When surveyed in the late 1970s, school psychologists reported administering projective measures of personality (e.g., House–Tree–Person, sentence completion) roughly twice as frequently as behavior rating scales Goh, Teslow, & Fuller, 1981). However, by the 1990s, the use of personality measures was shown to be on the decline, whereas use of rating scales and structured observations was steadily increasing (Hutton, Dubes, & Muir, 1992; Wilson & Reschly, 1996). Regarding observation in particular, the percentage of school psychologists reporting use of direct observation in at least 40% of their cases rose dramatically from 33% (Anderson, Cancelli, & Kratochwill, 1984) to 69% (Shapiro & Heick, 2004) over the last two decades of the 20th century. In fact, by 2008, over 95% of school psychologists reported that they had received training in systematic direct observation procedures, and over 90% indicated that they used direct observation procedures, to some degree, within their regular practice (Riley-Tillman, Chafouleas, Briesch, & Eckert, 2008). These findings suggest that direct observation has become both a core element of training and a mainstay of behavioral assessment in school-based practice.

ADVANTAGES OF DIRECT OBSERVATION IN SCHOOL-BASED ASSESSMENT

Although we have clearly seen an increase in the acceptability of observation procedures in recent decades (Riley-Tillman et al., 2008), the reality of the matter is that there are many different tools that school-based practitioners can use to assess student behavior (e.g., interviews, teacher rating scales, self-report measures). Furthermore, several of these tools (e.g., rating scales, Direct Behavior Rating) were designed to require minimal training and to be completed quickly by adults familiar with the student's behavior (e.g., parents and teachers). The question therefore inevitably arises of why we would choose to conduct

classroom observations, which are inarguably more costly in terms of time and resources? In some situations, the answer may simply be because it is required. For example, the most recent reauthorization of the Individuals with Disabilities Education Act (IDEA; 2004) requires that any student undergoing an initial or reevaluation to determine special education eligibility must be observed in the classroom. Although the law does not specify the type or quality of observation that must be conducted, it does note that "the public agency must ensure that the child is observed in the child's learning environment (including the regular classroom setting) to document the child's academic performance and behavior in the areas of difficulty" (Sec. 300.310). Observations may also be necessary as part of a functional behavioral assessment when conducting a manifestation determination to ascertain whether a student's behavioral problem is primarily the result of his or her disability. However, many other notable advantages of the use of direct observation are important to consider.

Directness of Assessment

One of the primary advantages of observation is the directness of the assessment. When we talk about the directness of assessment, we are describing how removed an assessment is in time and space from the actual occurrence of behavior (Cone, 1978). Highly direct assessment methods, such as direct observation, are those that can be used to assess behavior at the time and place it actually occurs. In order to assess a target behavior as it occurs, it is essential that the observer have a good definition of the behavior so that he or she can quickly distinguish between when it occurs and when it does not occur. This type of real-time recording is often more labor-intensive to carry out; however, it also results in data that are more accurate. On the other end of the continuum, indirect forms of assessment are those that are farther removed in time and space from when the behavior actually happened. Interviews and rating scales are examples of more indirect forms of assessment because they ask respondents to reflect on what behavior looked like at some time in the past. Retrospective judgments are often much easier to make, but they are also much more likely to be influenced by error because they rely on individuals' memories of events. Because respondents do not typically know what questions will be asked ahead of time, they are less likely to have paid focused attention to the specific target behaviors on which assessment is focused (imagine having to give a full report on the new restaurant in town when all you had been focused on was the food!).

As an example, consider if your doctor was interested in knowing the number of glasses of water that you drank in a given week. In order to obtain a highly direct measurement of water consumption, your doctor would first have to be explicit at the outset as to what constitutes a glass of water (i.e., 8 ounces = one glass). The doctor might then provide you with a notebook to keep in your pocket throughout the day and ask that you keep a running tally of the glasses as you consumed them. Although this system might interfere somewhat with your normal routines, you would likely have a fairly accurate record of water consumption by the end of the week. An alternative option would be for your doctor to simply ask you, upon entering his or her office, to estimate how many glasses of water you drank over the

past week. Certainly, this estimation would be much easier to do—probably requiring only a few seconds—but chances are that these numbers would also be far less precise. Use of direct assessment methods is therefore preferred in those situations in which more accurate data are needed, such as in the context of high-stakes decision making.

Flexibility

The inherent flexibility of direct observation is another advantage of the method. That is, observation can be used to assess a wide range of possible behaviors of interest across a number of different settings. Observation can be highly individualized, in that a code can be specifically tailored to focus on those behaviors of most interest within a given situation. As highlighted in Chapters 5–8, direct observation can be conducted across both classroom and non-classroom settings to assess behaviors ranging from student engagement to peer conflict. In fact, the only restriction in regard to what can be assessed using observation is that the behaviors of interest must be ones that can be seen and recorded.

Understanding Contextual Variables

An additional key advantage of direct observation is the ability to understand behavior within the setting in which it is a concern. As noted previously, prior to the 1970s, the traditional model of psychological assessment focused solely on understanding internal determinants of behavior, such as the traits an individual exhibited. It was believed that by understanding an individual's underlying personality structure, one could more reliably predict future behavior. Within a behavioral assessment model, however, behavior is not seen as existing in a vacuum but rather as being influenced by environmental factors. Unfortunately, although many available assessment methods are useful in generating a description of behavior, they contribute little toward understanding the impact of environmental events. Rating scales, for example, provide information with regard to the frequency with which particular behaviors occur in comparison to typical peers, but they do not provide additional information with regard to when or where the behaviors are more likely to occur. Although interviewing can be used to solicit information about environmental factors, such data are limited in that responses are filtered through the respondent's perceptions of the situation. That is, a classroom teacher may not have the cognitive resources to attend to the behavior of individual students when he or she is delivering instruction, or the teacher may interpret events differently than an objective observer (e.g., viewing yelling at a student as a form of punishment rather than as one of providing attention). Observation is therefore the only method that allows for the objective recording of environmental factors.

Provides Link to Intervention

Finally, because observations allow us to collect information concerning the context in which behavior occurs, observation data also have greater utility in suggesting next steps

for intervention. The results of rating scales, for example, help to verify that a problem exists (i.e., that there is a discrepancy between the student's behavior and that of his or her peers); however, they do not always provide information regarding why the problem is occurring or guidance regarding how the problem might be addressed. As is discussed in Chapter 8, observations can be designed to collect systematic information about potential triggers for problem behavior in the environment (i.e., antecedents), conditions under which problem behavior is more likely to occur (i.e., setting events), and events that may be maintaining problem behavior over time (i.e., consequences). This information can then be used to generate hypotheses regarding the function of problem behavior—that is, why the behavior is occurring.

USES OF DIRECT OBSERVATION IN SCHOOL-BASED ASSESSMENT

As we have tried to make clear, direct observation offers several advantages when compared with other behavioral assessment methods commonly used in schools. Although flexibility was stressed with regard to the ability to address a wide range of target behaviors, direct observation is also highly flexible in that it can be used for a variety of assessment purposes. The most common tasks to which direct observation is applied are (1) as one component of a multisource, multimethod assessment for classification purposes or to establish the presence of a problem; (2) to generate or test hypotheses as part of a functional behavior assessment; (3) for progress monitoring; and (4) for program evaluation. Direct observation has also produced outcome variables in studies evaluating the effects of psychosocial and medical interventions administered to large groups of children (e.g., Abikoff et al., 2004). This method therefore offers utility within both emerging problem-solving and traditional diagnostic models of service delivery.

Additionally, direct observation is unique in that it can be used to inform both nomothetic and idiographic decision making. Both classification and problem identification are nomothetic (i.e., interindividual) assessment tasks, meaning that comparisons are made to other individuals. In many of the assessments that you are likely familiar with (e.g., standardized testing, commercial rating scales), the target student's level of performance is compared with that of a large normative sample for the purpose of determining the target student's relative standing. This approach is often employed in assessment for classification purposes because it can be used to indicate the degree to which behaviors of concern depart from typical functioning. In contrast, idiographic (i.e., intraindividual) assessment focuses on the behavior of individuals without comparisons to others. Idiographic assessment tasks include the use of direct observation to inform the selection of academic or behavioral interventions and to monitor the effects of interventions on student behavior (e.g., Shapiro, 2004). Due to its sensitivity to environmental variables, direct observation has also been the primary assessment tool in functional behavioral assessment (e.g., Nelson, Roberts, & Smith, 1998; O'Neil, Albin, Storey, Horner, & Sprague, 2015).

Diagnostic and Eligibility Assessment

Direct observation data have long played an important role within the context of a traditional diagnostic model of service delivery. Within such a traditional model, students are referred for comprehensive assessment by a multidisciplinary team when their behavior is believed to deviate significantly from established norms or expectations. The goal of subsequent assessments is then to determine whether a student meets eligibility criteria for special education in order to inform a classification decision (i.e., eligible to receive special education services vs. not eligible). The decision of whether to enter a student into the special education system is undoubtedly one that involves high stakes. National statistics illustrate that a small percentage of students receiving special education services exit the system and that this is particularly true of students diagnosed with an emotional disability (SRI International, 2005). Furthermore, analysts have shown that the average incremental cost involved in educating students with disabilities in special education can be roughly $6,000 per student per year (American Institutes for Research, 2004). Given the high-stakes nature of the decisions being made, it is therefore critically important that more direct forms of assessment, such as observation, be used in order to either confirm or clarify problems reported through interviews, rating scales, and other similar tools.

When the goal of assessment is to make a diagnostic or classification decision, direct observation data can be used to provide evidence of either the level of impairment or the degree of discrepancy from peer performance. The IDEA definition of *emotional disturbance* (300.8c 4), for example, specifies that the disability involves:

> A condition exhibiting one or more of the following characteristics over a long period of time and to a marked degree that adversely affects a child's educational performance:
>
> (A) An inability to learn that cannot be explained by intellectual, sensory, or health factors.
> (B) An inability to build or maintain satisfactory interpersonal relationships with peers and teachers.
> (C) Inappropriate types of behavior or feelings under normal circumstances.
> (D) A general pervasive mood of unhappiness or depression.
> (E) A tendency to develop physical symptoms or fears associated with personal or school problems.

Using direct observation to collect quantitative data regarding the extent to which these characteristics are present (either in an absolute sense or in comparison to peers) can therefore help provide support that the behavior occurs to a "marked degree."

Similarly, diagnosis of *other health impairment* requires that the student demonstrate "limited strength, vitality or alertness, including a heightened alertness to environmental stimuli, that results in limited alertness with respect to the educational environment that is due to chronic or acute health problems . . . and adversely affects a child's educational performance" (300.8c9). Direct observation can therefore be used to provide an objective assessment of the level of strength, vitality, and/or alertness that a student demonstrates in the classroom environment.

Assessment within a Problem-Solving Model

Although the importance of data-based decision making within school settings has long been emphasized (e.g., Brison, 1967), the more recent accountability movement in education has brought data-based decision making to the forefront. For many years, the predominant approach to assessment in American schools was what has been labeled by some as a "wait-to-fail" model. That is, once an academic or behavioral problem exceeds a threshold of tolerability, the classroom teacher refers the student for a comprehensive evaluation to determine what supports are potentially needed to address the problem. The problem with such a model is that supports are not put into place until the problem has become sufficiently intense, making remediation even more difficult. What is needed is a more proactive model of assessment and intervention, in which students are provided with needed supports *before* problems become intractable.

In recent years, there has been increased focus on the use of a problem-solving model to address student problems. The problem-solving model emphasizes the use of data to develop and deliver student supports and consists of four sequential phases: problem identification, problem analysis, plan implementation, and plan effectiveness. Within the first phase, data are collected to verify the presence, and clarify the nature, of the problem. Assuming that a problem has been verified, the next goal is to determine why the problem is occurring and to use these data to inform the development of an intervention. After putting the intervention into place, the effectiveness of the intervention in improving student functioning is determined by monitoring student progress over time through the ongoing collection of data. Although several different types of data should be used to inform the problem-solving process (e.g., interviews, rating scales), here we wish to illustrate how direct observation data can be useful at each step along the way.

Clarifying the Nature of a Concern

The first step within the problem-solving process involves clarifying the nature of the concern. The goal here is to promote a common understanding among team members of what the potential problem looks like. In some cases, the person who has come forward with a concern may be too close to the situation to be able to describe the problem concretely or objectively (e.g., "He is constantly doing these little distracting things that are disrupting the classroom!"). Direct observation may therefore help to elucidate what the specific problem behaviors are and to what extent they occur. For example, a consultant (e.g., special education teacher, school psychologist) might enter the classroom and keep a running record of all of the disruptive behaviors exhibited by the target student. Following the observation period, the consultant and teacher could meet to review the information collected and establish concrete definitions for those behaviors deemed to be most problematic (e.g., occurring most frequently, interfering most significantly with instruction).

In other cases, it may be that the problem has been more definitively labeled, but the terms used to describe the concern are not sufficiently specific. As anyone who has ever been in a classroom will know, there are many ways to "not pay attention." Whereas one

student may stare out of the window or at a blank page in his notebook, another student's inattention may be much more disruptive in nature (e.g., talking with neighbors or turning around in her seat). In either case, observation by an external consultant or peer may be helpful in order to gain an objective understanding of the topography of the concern (e.g., What do these distracting behaviors look like? What does the student do when he or she is "not paying attention"?).

Finally, although a classroom teacher's concern may be narrowly focused on one specific problem behavior (e.g., task initiation), more typically referred students exhibit several problem behaviors in tandem. It therefore becomes necessary to prioritize which of several problem behaviors should be the focus of assessment and intervention efforts. It may seem that the easiest way to determine prioritization would be simply to ask the stakeholder which behavior he or she believes is the most important to target. One potential problem with this approach, however, is that the selected behavior may demonstrate high social validity but low practical utility. In other words, the teacher may think that the behavior is important to curb; however, reducing the behavior may not have any influence on meaningful outcomes. Consider, for example, a teacher who prioritizes pencil tapping as the primary problem behavior of interest. The repeated tapping of a pencil on a desk may feel like a form of slow torture to the teacher in the moment; however, intervening with this behavior is unlikely to lead to meaningful improvements in student achievement or well-being. External observers, on the other hand, may be better able to objectively identify those behaviors causing the greatest disturbance to either the learning of the target student or those peers around him or her. Observation may also be helpful in identifying those problem behaviors that occur early in a response chain (i.e., the sequence of linked behaviors), such that intervention efforts can be targeted at disrupting the chain before problem behaviors intensify.

Confirming That a Problem Exists

Once the nature of a concern has been clarified, the next essential step involves confirming that a problem does, in fact, exist. The existence of a problem has typically been conceptualized as a significant discrepancy between current and expected performance. Whereas objective standards often exist within academic domains for determining the level of expected performance (e.g., 20 digits correct per minute on a math computation probe), these determinations inevitably become more subjective when dealing with behavior. One way of defining the expected level of behavior is to simply ask the teacher about his or her expectations. For example, if Mr. Saunders believes that students should always remain in their seats unless they have received permission to get up, any instances of out-of-seat behavior would constitute a discrepancy from expected behavior. Collecting observation data to assess the actual number of times that a student is out of his seat then helps to clarify the scope of the problem and thereby inform the level of intervention that is needed. If observation data revealed that a student in Mr. Saunders' class typically got out of her seat once an hour, the discrepancy between current and expected behavior might be considered small enough to warrant minimal—or even no—intervention. If, on the other hand, the

student was found to leave her seat 10 times an hour, a more intensive intervention would likely be needed.

Alternatively, the expected level of behavior could be defined more objectively through the use of local norms. That is, observation data could be used to compare the current level of performance of a target student to the current level of performance of the target student's peers in the classroom. If results of observation validate that the target student's behavior is notably different from that of peers (e.g., the target student exhibits disruptive behavior during 50% of observed intervals, whereas typical peers exhibit disruptive behavior during only 5% of intervals), then such a finding would suggest the need for targeted intervention. If, on the other hand, the target student's behavior is found to be comparable to what behavior looks like in the larger classroom (e.g., the target student exhibits disruptive behavior during 50% of observed intervals and typical peers similarly exhibit disruptive behavior during 42% of intervals), the finding would suggest that resources should be directed instead toward the implementation of a classwide or group-level intervention (Batsche, Castillo, Dixon, & Forde, 2008).

Generating Hypotheses Regarding Why a Problem Behavior Is Occurring

Once a problem has been clarified and confirmed, the next phase in the problem-solving model involves generating hypotheses regarding why the problem may be happening. Simply knowing what the behavior looks like and how often it is occurring is not necessarily enough to inform intervention development. Rather, given that problems are often the result of an interaction between an individual and his or her environment, it is important to have an understanding of the situational/environmental variables at play. These variables also happen to be the ones that we have the greatest power to change (e.g., teacher response to inappropriate behavior, seating arrangement). As noted earlier in this chapter, one of the key strengths of direct observation is that it is possible to obtain objective information about both student behaviors and the context in which they occur. As is discussed in great detail in Chapter 8, systematically recording those events that both precede and follow instances of problem behavior can help to reveal patterns that may not have been otherwise obvious. These data can then be used to generate strategies for heading off problem behavior before it occurs as well as responding to both appropriate and inappropriate behaviors when exhibited.

Assessing Plan Effectiveness

After a plan has been put into place in the third phase of the problem-solving model, the final remaining task is to assess how well the plan worked. This assessment of plan effectiveness can be conducted using either formative or summative data. *Formative assessment* refers to the ongoing collection of data in order to determine progress over weeks or months. Although direct observation data have often been used within the context of single-case research studies to evaluate progress over time, we acknowledge the significant investment of time and resources that is needed to conduct ongoing observations in typical

school settings. Observation progress monitoring would therefore likely be used only with those students demonstrating the greatest levels of need or in those situations in which the stakes of decision making were extremely high. *Summative assessment,* on the other hand, refers to a one-time evaluation of learning or behavioral change. Direct observation data can be used within the context of summative assessment to determine whether the gap between expected and current performance has been reduced or eliminated by the end of an intervention period.

Use of Direct Observation across Assessment Purposes

Although observation data are ultimately used to serve different purposes within either a problem-solving or diagnostic model of service delivery, the data may be similarly useful in answering key assessment questions. For example, direct observation data can be used as part of a multimethod, multisource diagnostic assessment to confirm whether a problem identified through more subjective means (e.g., interviews, rating scales) exists as well as to illustrate the objective scope of the problem. Direct observation data can also be an essential piece of the puzzle when conducting a functional behavioral assessment to generate hypotheses as to why a problem behavior may be occurring and what should be done to address it (i.e., behavioral intervention plan) most effectively. Although later chapters do highlight observation codes that may be of particular utility in informing diagnostic decision making (e.g., codes examining behaviors characteristic of students with attention-deficit/hyperactivity disorder [ADHD]), we do maintain primary focus on the use of direct observation within a problem-solving model throughout this book given the important linkages to intervention development and monitoring.

CONCLUSIONS

Having discussed what direct observation is, the potential advantages of its use, and how it can be employed, one lingering question is who exactly should be empowered to take on the important role of observer? Although responsibility for student observations is often charged to student support personnel such as school psychologists, behavior specialists, and school social workers, the easy answer is that most school personnel are capable of conducting meaningful and informative school-based observations with the right training. We suspect that the reason why observations are not routinely conducted by a wider range of school personnel is simply that individuals do not know how best to observe or what to do with the resulting information. If answers to these questions are not clear, the likely result is both discomfort and wasted time.

As William Penn (1909–1914) wrote, "Time is what we want most, but . . . what we use worst." We know that time is a precious commodity, especially in schools, and the goal of this book is therefore to help you to make the most of your observation time. As we have outlined in this chapter, there are several advantages to collecting observation data as part of a school-based assessment. However, if you were going to commit the time and resources

to conducting an observation, you would want to know that the data you collect would have the greatest utility and defensibility possible. Over the next nine chapters, we aim to provide you with both the knowledge and tools needed to conduct meaningful observations of a wide range of student behaviors.

This book is divided into three main sections. The first section, *Foundations of Direct Observation,* is intended to provide you with the basic conceptual foundations of direct observation. Across the next three chapters, we review the basic considerations involved in conducting school-based observations (Chapter 2), describe the most important indicators of high-quality observation methods (Chapter 3), and suggest guidelines for maximizing the reliability and validity of decisions based on observation data (Chapter 4). The second section, *Use of Specific Observational Codes,* moves from a general discussion of the overall assessment method to a more specific discussion of extant codes that are available for specific purposes. Four chapters introduce a number of evidence-based observation codes that were designed to assess student behavior in classroom settings (Chapter 5), the classroom environment (Chapter 6), student behavior in non-classroom settings (Chapter 7), and both student behavior and the environmental context within functional assessment (Chapter 8). In those cases in which an extant, validated observational code does not fit the target behavior(s) of interest, however, it may be necessary to create a new code. We therefore provide general guidelines for developing a unique code in Chapter 9, as well as a library of operational definitions and a range of sample coding forms within the appendices. The third and final section, *Using Assessment Data to Inform Decision Making and Intervention,* consists of one chapter focused on what to do once observational data have been collected. Specifically, this chapter presents strategies for summarizing, graphing, and interpreting direct observation data across different purposes of assessment.