Direct Behavioral Observation in School Settings: Bringing Science to Practice

Matthew K. Nock, *Harvard University*
Steven M. S. Kurtz, *New York University Child Study Center*

Schools provide a useful, controlled setting for evaluating child behavior problems, yet direct observational coding procedures evaluated by child researchers have not been widely incorporated by practicing clinicians. This article provides a summary of procedures useful to clinicians performing direct behavioral observation in school settings. We describe the need for and usefulness of comprehensive school observations; provide a primer on the identification, definition, and assessment of target behaviors; and outline and discuss specific clinical procedures, including formulating primary referral questions, interviewing teachers, describing the classroom context, and conducting the observation. We also provide practical advice for synthesizing the obtained information into a report that guides clinical intervention. A sample of school observation coding forms and guidelines for report writing are also included to facilitate the use of these techniques by clinicians and teachers involved with the child.

Approximately 20% of school-aged children and adolescents (9 to 17 years old) suffer from a current diagnosable mental disorder (Shaffer et al., 1996; U.S. Department of Health and Human Services, 1999). Rates of specific disorders vary, and this overall figure includes a wide range of potential problems, including difficulties with attention and concentration, aggressive or oppositional behavior, anxious or depressive behavior, substance use, and learning and mental disabilities. Given that virtually all children in the U.S. and other industrialized countries are required to attend school on a consistent basis, many of these behavior problems manifest in school settings and have a significant, negative impact on child development and social and academic functioning (Barkley, DuPaul, & McMurray, 1990; Cole, 1990; Kandel & Davies, 1986; Lambert & Sandoval, 1980). Psychologists and other mental health professionals are routinely asked to provide brief or, in some cases, ongoing consultation in school settings for a wide range of child behavior problems. This often takes the form of a *school* or *classroom observation*, which refers to measurement procedures in which child behaviors in the school or classroom are systematically monitored, described, classified, and analyzed, with particular attention typically given to the antecedent and consequent events involved in the performance and maintenance of such behaviors. In fact, the recent reauthorization of the Individuals With Disabilities Education Act (2004) makes the performance of a school observation a required component of the initial evaluation and re-evaluation for students with behavior problems, and includes additional recommendations for the performance of functional behavioral assessments in school settings.

Although many of the aforementioned behavior problems manifest in multiple contexts (e.g., home, community), the school setting provides a structured environment in which the clinician is able to evaluate the child's behavior across multiple behavioral domains. For instance, the school setting provides a contained environment in which the clinician can observe and evaluate the ongoing, bidirectional interactions between the child and the peers and adults present in the school setting. In addition, schools enable the clinician to assess the child's ability to engage in and master various academic and developmental tasks. School settings also require the child to perform in both structured and unstructured contexts. Each of these opportunities may exist to some degree in a clinical or home-based assessment; however, schools are perhaps the most feasible and ecologically valid context in which to collect these data.

Despite the high prevalence of child behavior problems and the legal requirements necessitating the performance of school observations, standard procedures for collecting such data do not exist. Although significant research efforts have evaluated various formal assessment procedures for such purposes, these methods have not been widely adopted by practicing clinicians. This is perhaps due to differences in training and resources between research and clinical settings. For instance, Abikoff and colleagues (1977, 1980) used observers trained for approximately 50 hours, who coded child behavior for an average of 96 minutes across six observation sessions in order to obtain data to reliably differentiate hyperactive from nonhyperactive children. This distinction would be even more difficult to make in nonresearch settings given the lack of normative data on the behaviors associated with each diagnostic category. Moreover, in addition to

Cognitive and Behavioral Practice 12, 359–370, 2005
1077-7229/05/359–370$1.00/0
Copyright © 2005 by Association for Advancement of Behavior Therapy. All rights of reproduction in any form reserved.

diagnostic information, clinicians also need to identify the functional determinants of the problem behaviors, as well as the varying characteristics of the child, teachers, peers, and classroom environment that may be useful for designing and modifying interventions aimed at remediation of the problem behaviors. Procedures for performing these tasks are not neatly encapsulated in any existing research protocol available to practicing clinicians.

In an effort to help fill this science-practice gap, the purpose of this article is to provide a summary of procedures useful to clinicians performing behavioral observations with school-aged children. Rather than presenting a detailed description of any one coding system currently used by researchers in this area, we describe general guidelines and procedures for the implementation of a comprehensive, idiographic school observation. In doing so we draw directly from our work developing and testing methods for conducting behavioral observations in schools. After discussing the advantages and disadvantages of direct observation procedures, we provide a primer on the identification, definition, and assessment of child behaviors to increase the relevance of this information for those who may be unfamiliar with behavior observation techniques. Next, we outline and discuss specific clinical procedures, including formulating primary referral questions, interviewing the teachers, describing the classroom context, and conducting the observation. Finally, we provide practical advice for synthesizing the obtained information into a report that facilitates clinical intervention by clinicians, teachers, and parents involved with the child.

Advantages and Disadvantages of Direct Observation Procedures

School observations offer several advantages over other assessment methods, and these advantages can greatly enhance the effectiveness and efficiency of clinical interventions. First, school observations provide a means of objective measurement of a wide range of behaviors as they occur in the natural environment. Indeed, observational data are free from most sources of bias associated with behavior rating scales such as social desirability, mood of the rater, rater attributions about culture or socioeconomic status, and halo effects (e.g., Abikoff, Courtney, Pelham, & Koplewicz, 1993; Cohen & Kasen, 1999; Edelbrock, 1988) and have been shown to better predict future adjustment than parent or teacher reports of child behavior (e.g., Patterson & Forgatch, 1995). In addition, the use of direct observation allows for the description and evaluation of specific behaviors. Even the most comprehensive behavior rating scales do not contain items representing the full range of possible behavior problems exhibited in the school/classroom context. Also, behavioral descriptions presented in rating scales typically are vague and more likely to introduce criterion error variance, compared with the use of direct observation methods that allow for a more precise evaluation of the target behaviors.

Second, observational data have greater external or ecological validity than behavior rating scales, as they provide a measurement of the behavior as it is actually occurring in the school context. Third, while behavior rating scales provide information about the frequency and severity of behavior problems, observational data also provide information about the *functions* of such behaviors (Hanley, Iwata, & McCord, 2003; Kazdin, 2001; Miltenberger, 1997).

That is, behavioral observation can provide information about the purpose or cause of a given behavior by evaluating the antecedent and consequent events that maintain the target behaviors. Although rating scales designed to gather information about antecedent and consequent influences on behavior have been developed (Durand & Crimmins, 1988; Nock & Prinstein, 2004), a true functional behavioral assessment generally is not possible without directly observing the child's behavior. The ability to observe child behavior in a real, structured setting represents an important advantage of school observation over other assessment methods.

Despite the advantages of using direct observation methods, several disadvantages of this approach should be noted. First, the performance of a school observation is more costly in terms of time, money, and resources than other measurement methods such as parent-, teacher-, or child behavior rating scales. School observations typically require at least 1 hour of direct observation in addition to time and expenses for travel, preparation, and report writing, compared to the relatively little time required to complete behavior rating scales. Clinicians should consider offering a behavioral observation in the school setting as part of initial or ongoing assessment services or as part of a larger psychological evaluation. Moreover, school observations require training and utilization of mental health professionals, while most rating scales can be administered and scored by nonclinical personnel. Second, the identification and evaluation of a specific behavior during a discrete time period may provide limited data about a child's behavior in the classroom or in other settings. That is, although data from teacher ratings tend to be more vague and global than data from direct observations, the former provide information about a wide range of behaviors from the entire school day. In contrast, direct observation techniques provide a narrower sample of specific behaviors. Third, there are potential methodological problems with the use of direct observation, such as child or teacher reactivity to the presence of the observer, observer drift, and perceptual biases of the observer (Kazdin, 1978; Skinner, Dittmer, & Howell, 2000).

These advantages and disadvantages should be weighed

when deciding whether a school observation is appropriate for a specific child. To be sure, all children referred for clinical services will not require a school observation. Although school observations could always provide potentially useful clinical information, given the time and resources required and the potential limitations of this approach, they should generally be reserved for cases in which the form and/or function of a behavior problem present in the school is unclear. We recommend a *stepped model of assessment*, in which the intensive evaluation of specific behavior problems should be implemented only after other more global and less intensive assessment methods have been exhausted. This approach differs from stepped care models of psychotherapy (e.g., Haaga, 2000), given its focus on assessment rather than intervention, and differs from multiple gating screening techniques (e.g., Loeber, 1990; van Lier, Verhulst, & Crijnen, 2003), given its focus on an individual child rather than a mass screening approach. For instance, a child with self-, parent-, and teacher-reported symptoms of obsessive-compulsive disorder occurring consistently across multiple contexts will not require a school observation because such an assessment is not likely to add unique information. Alternatively, in a case of a child referred for aggressive and noncompliant behavior with some teachers but not others, a school observation may be warranted as it could provide objective, ecologically valid information about the form and function of the target behaviors that may not otherwise be accurately evaluated.

Structured School Observation Coding Systems

A number of structured systems for coding direct observation data collected in school settings have been developed and evaluated over the past several decades and may be of interest to the clinician conducting a school observation. These systems can be used by the clinician when they match the problem behavior and setting in question. Alternatively, in situations where there is not a direct match between the system and the identified problem, these systems may serve more generally as a methodological guide for structuring or implementing a more individualized school observation. One of the earliest and most thoroughly evaluated systems, the Classroom Observation Code (COC; Abikoff & Gittelman, 1985), uses a time sampling procedure to classify child behaviors according to 14 observational categories. As suggested by the different behavioral categories (e.g., interference, solicitation, off-task, physical and threats of aggression, noncompliance, motor movement, extended verbalization, daydreaming), the COC has been used primarily with hyperactive, inattentive, and disruptive children. The COC has adequate interobserver reliability, discriminates between hyperactive and nonhyperactive children, and has no detectable observer effect on child behavior (Abikoff et al., 1977, 1980). The application of this coding system has extended beyond mainstream classrooms, and its interobserver reliability and concurrent and discriminative validity have also been supported in classroom settings within psychiatric hospitals (Horn, Conners, Wells, & Shaw, 1986). Among the limitations of these findings, however, the reliability and validity estimates generated were based on an average of 11.6 days of observations for each child (Horn et al., 1986), and this system has limited generality beyond hyperactive behaviors in children.

Another direct observation coding system that has received considerable empirical support is the School Observation Coding System (SOCS; McNeil et al., 1991) and the subsequent Revised Edition of the School Observation Coding System (REDSOCS; Jacobs et al., 2000).[1] The SOCS and REDSOCS both use a modified time sampling procedure and classify behaviors on three behavioral domains: (1) appropriate versus inappropriate behaviors, (2) compliant versus noncompliant behavior, and (3) on-task versus off-task behavior. The behavioral categories are broader than those used in the COC, and the developers of the SOCS and REDSOCS report less training time required for the observer to reach performance criterion. Similar to the COC, the SOCS and REDSOCS have been used primarily with children with disruptive behavior disorders, and studies using these measures have supported the interobserver reliability and the concurrent and discriminative validity of these measures (Jacobs et al., 2000; McNeil et al., 1991). Although the structured nature of the REDSOCS is among its strengths, this system too is limited in its ability to assess idiosyncratic child behaviors that may occur in school settings.

Several direct observation coding systems are also offered as adjunctive components of widely used, multirater assessment packages. For example, the Child Behavior Checklist (CBCL; Achenbach, 1986) includes a Direct Observation Form (DOF) to accompany teacher-, parent-, and self-report forms. Similarly, the Behavior Assessment System for Children (BASC; Reynolds & Kamphaus, 1998) includes a Student Observation System (SOS) to accompany teacher-, parent-, and self-report forms. Both of these coding systems are easy to learn and implement, requiring minimal training, are readily available as part of widely used assessment packages, and sample a wide range of behavior problems (96 on the DOF and 65 on the SOS) and adaptive behaviors in a short time period (10 to 15 minutes of direct observation for each administration). However, the use of a list of predetermined behavior problems, a small time sampling period, and minimal observer

[1] It is our belief that a more appropriate name for this measure is the Youth Assessment of Needs for Kids Exhibiting Emotional problems in School (YANKEES). However, we acknowledge a potential bias, as both authors are from the New York metropolitan area.

training, as well as a failure to consider functional determinants of observed behaviors, limit the information derived from the use of these measures.

Clinicians should consider what type of behavioral information is desired from the observation and what resources are available when deciding what type of direct observation coding system to utilize in any given case. In instances where one of these systems is not appropriate (e.g., target behavior is not assessed sufficiently by the system), the clinician should conduct an assessment grounded in the principles of behavioral assessment (see Kazdin, 2001; Miltenberger, 1997; Shapiro & Kratochwill, 2000). We have written previously about adapting structured, evidence-based psychosocial treatments for idiographic clinical use guided by psychological science and behavioral assessment (Nock, Goldman, Wang, & Albano, 2004), and offer here a similar approach focused more directly on conducting school observations. We first offer a primer on behavioral assessment as it applies to children in school settings for those unfamiliar with such methods and their application with this population and in this setting.

A Primer on Identifying, Defining, and Assessing Child Behavior

Identify the Problem Behaviors Via a Thorough Clinical Interview

The valid assessment of child behavior problems requires the collection of information that spans multiple domains of functioning, informants, measurement methods, and environmental contexts (Campbell & Fiske, 1959; Shapiro & Kratochwill, 2000).

Multiple domains. Families often present for assessment with one identified "chief complaint"; however, this is not necessarily the only area requiring intervention or the most ideal primary treatment target. Rather than proceeding immediately with intervention, it is important to assess multiple domains of functioning in order to identify all potential areas of maladaptive and adaptive functioning and to collect information about the structure and function of each identified behavior.

Multiple informants and contexts. Decades of research have demonstrated that interinformant agreement on the presence and severity of child behavior problems is quite poor across various domains of functioning (e.g., Achenbach, McConaughy, & Howell, 1987). There are many factors contributing to this variability, including situational specificity of behavior, informant bias, method variance, and type of behavior problem. For example, agreement across all informants (child, peer, teacher, and parent) is generally better for externalizing compared to internalizing problems (Edelbrock et al., 1986; Ledingham, Younger, Schwartzman, & Bergeron, 1982). In addition, internalizing problems such as depressed mood, anxiety, and somatic complaints are reported more frequently by children than by their parents (Ivens & Rehm, 1988; Kashani, Orvaschel, Burk, & Reid, 1985). Nevertheless, both parent- and child-report of child behavior problems have been shown to be significantly related to measures of external validation (i.e., service utilization, evidence of behavioral impairment, etc.; see Jensen et al., 1996). Taken together, these findings suggest that despite low levels of agreement across informants, multiple informants should be used in child assessment given the unique information each provides about child behavior in different contexts and from different perspectives.

Multiple methods. Self-report procedures are typically characterized by higher sensitivity for detecting constructs of interest (e.g., Prinstein, Nock, Spirito, & Grapentine, 2001), perhaps due to willingness of children and adolescents to endorse items on self-report they might be hesitant to discuss in interview-based assessment. In contrast, interview-based approaches generally allow for greater specificity through follow-up questions. Given differential responding for various measurement methods, clinicians should employ multiple assessment methods whenever possible in order to ensure accurate identification and diagnosis of clinical behavior problems.

Define the Target Behaviors

Once a problem area has been identified, the target behaviors must be defined in a manner that is *observable*, *measurable*, and *specific*. The criterion of being *observable* refers to the definition of the target behaviors in terms that are readily apparent to more than one individual. For instance, "the number of times Johnnie kicks a classmate" satisfies the observable criterion, while "the number of times Johnnie gets angry" does not, given that "anger" refers to an affective state that is not reliably apparent between independent observers. If a behavior is conceptualized in a way that makes it observable, a *measurement* strategy can be devised to quantify its occurrence. *Specificity* refers to the precision of the defined target behaviors, such that boundary rules are established to indicate when a behavior has or has not occurred. Consider this definition: "Kicking a classmate refers to making physical contact between Johnnie's foot and any part of another student but does not include instances in which Johnnie stops his foot before contact is made." This specific definition precisely indicates when the target behaviors have occurred and when they have not.

Although undesirable or maladaptive behaviors are often the focus of assessment, as in the examples noted previously, the goals of clinical intervention should center on developing *replacement* or *alternative* behaviors (Blader, Nissen, Fleiss, & Kurtz, 2000). *Replacement behaviors* refer to the ultimately desired, adaptive behaviors that are

incompatible with the problem behavior. For instance, if the problem behavior is "punching a teacher in response to a difficult task demand," a replacement behavior might be "complying with difficult task demands while keeping hands and feet to self."

Given it is typically not expected that a child will change directly from a problem behavior to an adaptive, incompatible behavior, it often is helpful to first develop *alternative behaviors*, which refer to acceptable behaviors whose performance decreases the probability the problem behavior will occur. Alternative behaviors are not necessarily incompatible with the problem behavior, and may be approximations of the desired behavior. If the alternative behavior is functionally equivalent with the problem behavior, this may increase the probability of maintenance of the alternative behavior. However, functional equivalence is also not a necessary criterion for the development of alternative behaviors. For instance, in the example above an alternative behavior might be "requesting a pass to go to the guidance office in response to difficult task demands." If the function of the child's problem behavior is determined to be avoidance of the task demand, and this alternative behavior also functions to avoid the task demand, then the two are considered functionally equivalent. Although not the ultimately desired outcome, leaving the room is a more acceptable behavior and considered a step closer to the desired outcome. Another alternative behavior might be "keeping hands to self in response to a difficult task demand." In this case, the child should be rewarded if he does not hit the teacher in response to a difficult task demand, even if he performs some other undesirable behavior such as yelling at her. Although yelling at the teacher is of course not the ultimately desired behavior, and "keeping hands to self" is not functionally equivalent with avoiding the task demand, it is considered a closer approximation of the desired behavior than is the problem behavior and thus is considered a favorable alternative behavior.

Define Events Related to the Target Behaviors

Given that myriad factors in a child's environment, internal and external, influence the occurrence of given behaviors, school observations will be most effective when contextual influences are identified and measured in a systematic manner. As a general rule, in addition to defining and measuring the target behaviors, the observer should also provide a description of antecedent and consequent events associated with each target behavior.

The observation should consider different types of antecedent influences, including information about potential *establishing operations* or *setting events*, *discriminative stimuli*, and *prompts*. *Establishing operations* and *setting events* refer to factors that alter the value of reinforcers and the likelihood of engaging in specific behaviors to obtain reinforcement. For instance, time of day, exhaustion, and hunger are all establishing operations that might influence child behavior (e.g., children may be more likely to perform behaviors reinforced with access to a stimulating toy when other toys are unavailable rather than abundant). *Discriminative stimuli* refer to events that indicate a particular behavior is likely to be reinforced. For example, some behaviors may only be reinforced or punished in the presence of a specific teacher or staff member but not others. *Prompts* refer to events that directly guide or facilitate task performance, such as instructions, commands, or reminders from a teacher.

The observation should also consider different types of consequent events, including information about the presentation or removal of different events that follow the target behaviors. The target behaviors may be followed by the presentation of a desired event (i.e., *positive reinforcement*), such as attention from the teacher or laughter by peers; the removal of an aversive event (i.e., *negative reinforcement*), such as the removal of the teacher's demands to go to the principal's office; or by the lack of any event, such as being ignored by the child's teacher and peers.

Information about such antecedent and consequent events is critical in understanding the determinants of the target behaviors and is useful in treatment planning. Without such data one is left without the ability to see behavioral patterns and inevitably will either make no objective cause-and-effect hypotheses or make ones that are (a) without an empirical basis, (b) likely to be less accurate and less parsimonious, and (c)lead to less efficient interventions.

Assessment Strategies

Once the target behaviors have been adequately defined, the observer must monitor the child to record the occurrence of the behaviors, as well as the presence of antecedent and consequent events. This is in contrast to behavior rating scales, in which child, peer, teacher, or parent retrospective report is used to describe the prior occurrence of the target behaviors. There are several different methods for recording the occurrence of the target behaviors in school observations. These methods can be used alone, in sequence, or in combination.

Descriptive method. One way the observer can record the occurrence of the target behaviors is simply by providing a specific description of the behavior as it is performed. In doing so, the observer should provide an objective account of the mechanics of the behavior, the amount of time the behavior is performed, and the intensity of the movements. As mentioned, the observer should also describe the antecedent and consequent events surrounding each performance of the target behavior.

This descriptive method typically should last at least 10 to 15 minutes, depending on the frequency of the target behaviors, and is often most useful when the specific char-

acteristics and potential determinants of the target behaviors are not yet known. An example of a coding sheet following a descriptive format is presented in Figure 1. In many cases, it is advisable to follow the descriptive method with the use of a more structured, time-sampling technique, such as one of those described below, to generate more systematic data once the potential contingencies have been operationalized during the initial observation period.

Checklist method. Another method for recording the occurrence of the target behaviors is the use of checklists in which the observer makes check marks to record various aspects of the target behavior. Behaviors typically are listed on one axis of a recording form and time periods or classrooms are listed on the other axis. There are several different types of checklists commonly used for direct behavioral observation.

Behavior checklists refer to simple forms in which the observer uses check marks to indicate which behaviors from a long list of potential behaviors occurred during the observation period. For instance, the BASC-SOS (Reynolds & Kamphaus, 1998) contains a section with a list of 14 behavioral categories from which the observer indicates whether a behavior from each category occurred during a 15-minute observation period. A variation on the behavior checklist is *frequency recording*, in which the observer indicates how often the behavior occurred using a check mark for each instance of the behavior. An example of a frequency recording checklist is presented in Figure 2.

Interval recording refers to a technique in which the observer indicates (again using check marks) whether the target behaviors occur during a specific time interval. Interval length varies depending on the frequency of the behavior, the amount of time allowed for the observation, and the skill of the observer in monitoring and recording child behavior. As an example, the BASC-SOS includes an interval recording section with the following instructions: "At the end of each 30-second interval, observe the child's behavior for approximately 3 seconds (for example, when the stopwatch reads 0:30–0:33). Then place a check mark in the time column next to each category of behavior that occurred during that interval." This procedure is performed repeatedly across a 15-minute time period and provides a contemporaneous assessment of behavior as it occurs. The percentage of intervals during which a given behavior occurred can be calculated to provide information about the frequency of adaptive and maladaptive behaviors. An example of an interval recording form is presented in Figure 3.

As is the case with descriptive methods, checklists can also be used to record information about antecedents and consequences of target behaviors. This is performed most accurately using interval recording techniques given the multiple variations of antecedent-behavior-consequence that are possible, which can be difficult to report retrospectively. The prespecified behaviors coded on check lists can be general (e.g., physical aggression) or specific (e.g., kicks others, hits others, throws object, etc.). The specificity of behavioral coding will depend on the needs and abilities of the observer and will likely vary across behaviors, settings, and children. The strength of this approach lies in the ability of the observer to generate data about the *conditional probability* of different target behaviors or consequences. Conditional probability refers to the likelihood of a behavior given specific settings or situations. Through the assessment the observer obtains information about the probability of each target behavior in the presence of specific antecedents (e.g., presence of teacher, use of prompts, occurrence of different activities) and consequences (e.g., positive vs. negative vs. ignoring). For instance, Figure 4 depicts a sample coding form used to record antecedent-behavior-consequence relations in a specific classroom period during 16 time sampling intervals (i.e., an 8-minute period with coding every 30 seconds). The aggregation of data from this period shows that in terms of consequences, physical aggression was followed by a positive response from peers, in this case attention and laughter, during 100% (8/8) of the intervals, and was followed by a negative consequence from the teacher only 12.5% (1/8) of the time. In contrast, positive social interactions were followed by no observable teacher (0%) or peer (0%) initiated consequences. These data suggest that the child's physical aggression may be maintained by the positive reactions it produces from peers, and that there are no positive consequences for engaging in positive social interactions. These data can be analyzed and interpreted according to antecedent, behavior, or consequence based on the observer's hypotheses about what factors may be initiating or maintaining the target behaviors. It is critically important in the formulation of the hypotheses and the case conceptualization to consider the conditions under which the target behaviors *do* and *do not* occur. It is the rule, rather than the exception, that we observe marked intersituational variability in disruptive behaviors.

Clinical Procedures

Formulating the Primary Referral Question

After conducting a comprehensive initial evaluation with the child and family, the clinician should formulate the primary question to be answered by the school observation. The specification of the primary referral question is especially important given the range of observational and coding techniques available to the clinician. For instance, if the question relates to whether and in what classroom a specific behavior is being performed, the

Situation	Behavior	Antecedents	Consequences
Sitting in circle time 9:00 am	Slapped peer sitting immediately to his left with open hand	None observed	Peer and teacher both ignored
9:03 am	Yelled "NO!" at teacher and remained seated	Teacher gave specific command for all students to return to their seats	Teacher ignored, students laughed
9:05 am	Got up and sat in seat	Peer came over to him and whispered in his ear	Teacher gave specific, labeled praise to target child

Figure 1. Sample Descriptive Coding Form.

Behavior	Frequency
Fidgeting in seat	IIII
Getting out of seat	III
Running around classroom	I
Interrupting others	IIII
Physical aggression toward peers	I
Physical aggression toward teachers	
Verbal threats of aggression toward peers	II
Verbal threat of aggression toward teachers	I

Figure 2. Sample Checklist for Coding Child Behavior.

Behaviors	30"	1'	30"	2'	30"	3'	30"	4'	30"	5'	30"	6'	30"	7'	30"	8'	30"	9'	30"	10'
Inappropriate Movement	√				√															
Inattention			√	√					√	√						√	√	√	√	√
Physical Aggression	√	√																		
Self-Injurious Behavior						√	√							√						
(continued)																				

Figure 3. Sample Interval Coding Form.

Target Behavior	Antecedents	Consequences						Comments
		Teacher +	Teacher -	Teacher 0	Peer +	Peer -	Peer 0	
Physical Aggression	Teacher left room Child took his toy None observed		I	IIIIII I	IIIIII I I			
Positive Social Interaction	Teacher specific prompt None observed			IIIII III			IIIII III	
(continued)								

Figure 4. Sample Interval Coding Form With Antecedents, Behaviors, and Consequences.

observer may consider using a frequency recording technique across several different classrooms. In contrast, if the question is focused on determining the function of a given behavior with the goal of providing information for the treating clinician, an observational method that includes information about the antecedents and consequences of the target behaviors is indicated. Included in the primary referral question should be a clear statement of the target behaviors, as well as a statement of the specific goals of the assessment.

Collaborating With School Psychologists and Teachers

School psychologists often spend a large amount of time developing and implementing behavioral assessment and intervention programs in school settings. If there is a school psychologist at the target child's school, he or she is likely to be the clinician conducting the behavioral observation and working with the teacher(s) and/or outside clinician on reporting on the results of the assessment and developing a treatment plan. In instances where a school psychologist is not present or is unable to participate in the behavioral assessment, the outside clinician may need to visit the school him- or herself to conduct the assessment. In either case, every effort possible should be made to work openly and collaboratively with the mental health professionals and teachers involved with the child.

Prior to conducting the observation, the observer should contact the classroom teachers to obtain permission to perform the observation and to begin to gather information about the target behaviors. Most of the necessary information typically can be collected in a brief meeting or telephone interview. There are several domains from which the observer should attempt to gain information. First, the observer should inquire about the specific target behaviors. Whether or not the teachers have already provided information about specific behavior problems as part of the initial comprehensive evaluation, the observer should obtain a clear and complete description of the target behaviors from the teachers' perspective. This should include maladaptive behaviors as well as adaptive/desired behaviors. Second, the observer should inquire about the presence of any academic difficulties.

Third, the observer should generate a list of potential contextual influences of the target behaviors. Questions for the teachers should include: When/where/with whom does this behavior most commonly occur? In addition: Have you noticed what is typically happening before this behavior occurs? And: What do you do, or what do the other children do, after this behavior occurs? Critically, one also must solicit the same types of information from the teacher about the situations in which the target behaviors do not occur. This list of antecedent and consequent events provides useful information for generating hypotheses and conducting the actual behavioral observation. The answers to these questions should be taken into consideration when planning the observation. If the behavior occurs infrequently, or is a "low base-rate" behavior, the observer should plan to monitor the child's behavior only when the behavior is most likely to occur. For instance, if the child engages in physical aggression, but seems to do so only when engaged in outdoor physical activities, it would be advisable to plan the observation session during a period that includes outdoor physical activities. In cases when it is not possible to schedule the school observation during a situation that seems to influence the occurrence of the target behaviors, the observer may want to work with the teacher to architect the setting events or situations that are likely to produce the target behaviors in order to monitor the relations among antecedents, behaviors, and consequences (e.g., scheduling an outdoor activity while the observer is present, rearranging classroom activities or interactions in order to increase the probability of the target behaviors).

Finally, the observer should inquire about what types of interventions are currently in place in the classroom. It would also be helpful to know what interventions have been tried in the past and what has or has not worked in modifying the target behaviors. In some cases it is advisable to ask the teacher to implement a current or past intervention that has or has not been effective in order to understand which factors influence the child's behavior or which aspects of the intervention content or delivery might require modification. Specific parent consent is advised if choosing to precipitate the undesired behavior.

Describing the School/Classroom Context

Once in the child's classroom, the first task is to make contact with the teacher and have him or her identify the child to be observed and to select a location from which to observe the child's behavior. It is best to do this in a way that limits the probability of child reactivity to the observation. The child should not know the observer is visiting the classroom for the purpose of observing his or her behavior, as doing so would compromise the validity of the observation. The position of the observer in the room should allow for an unobstructed view of the child while limiting the probability of child reactivity to the observation (i.e., out of sight of the child). The back of the classroom usually is best for these purposes.

Next, the observer should spend a few minutes collecting information about the contextual characteristics of the classroom, with particular consideration to aspects of the classroom that may affect child behavior. For instance: Is the classroom seating and furniture arranged in a way that limits distracting stimuli, background noise, outside class interruptions? Are traffic patterns safe and well-defined? Is the placement of the child in the class-

room conducive to good attention and consequating by the teacher? How many students and teachers/aides are in the classroom? Are the class rules and consequences, as well as the daily schedule of activities, posted in a visible location? How large is the room, how many windows and doors are there and where are they located? Each of these factors can affect child behavior and may be of interest to the clinician.

Observing the Child's Behavior

After the classroom environment has been described, the full attention of the observer should shift to monitoring and coding the child's behavior. The first task of the observer is to select observation techniques from those described above. This should be determined prior to the actual school observation, and the selection of a coding method is contingent upon the goals of the observation.

The actual school observation typically lasts 1 to 2 hours during which the primary goal of the observer is to record data regarding the planned target behaviors. After an initial coding period, the observer should review the data and decide to monitor and code any new behaviors noted during observation period. At the conclusion of the observation period, the observer should speak with the teacher to determine the typicality of the child's behavior during the observation period, noting the extent to which these results represent an average day for the child.

Although studies of the reliability and validity of behavioral coding systems have compared ratings of groups of children referred for behavior problems with those of nonreferred children, we do not believe it is necessary to record the behavior of a "comparison child" when conducting a school observation for an individual child. Indeed, the extra data collected will not allow the clinician to draw valid inferences about any obtained differences given potential problems with sampling error, selection bias, and natural variability in the children's behavior. We believe it is more clinically useful to collect data on the conditional probability of the behavior of the target child.

Preparing the Report

When composing a school observation report, it is important to know your audience. Consider who will be reading the report and who will be using the specific recommendations generated. Regardless of whether the report is ultimately intended for use by clinical or school staff, it will likely be read by the parents of the target child and thus should be written in a way that is easily understood by the lay person.

Format of the Report

There are many ways to structure individual reports from school observations. Regardless of how the sections are organized, all important information must be contained in the written report. A clinician can spend hours preparing, observing, and analyzing data, but if the information is not synthesized and presented in the written document, it will not have an effect on the child's school environment. An outline of suggested major areas and guiding questions for the school observation report is presented in Table 1.

Incorporating the Data

One of the key strengths of school observations is that they are based on direct, objective data. The clinician should capitalize on this strength when preparing the report. The data should be presented in the report so the reader (whether clinician, teacher, or parent) can see the evidence supporting the interpretations made by the observer. For instance, when describing the determinants of a given behavior, the observer should indicate how often the behavior occurred, as well as how often different events preceded and followed the behavior, thus drawing evidence-based conclusions about potential determinants of the target behavior. Consider two different hypotheses about the factors maintaining a child's behavior, one based on the clinician's unmeasured beliefs about the nature of the child's behavior:

> Alex often gets out of his seat and does not follow teacher commands. This is likely due to his lack of respect of authority figures and his inability to follow instructions.

and one based on the data collected in the school observation:

> Despite remaining in his seat and maintaining appropriate behavioral control during 77% of the observed intervals, this behavior was not followed by a positive consequence on any occasion. In contrast, Alex's inappropriate movement was followed by positive peer interactions 50% of the time (e.g., eye contact, smiles, and giggles from peers), nonspecific teacher reprimand 10% of the time (e.g., "Alex!"), and was ignored 40% of the time. Taken together, these results suggest Alex's appropriate classroom movements are not being rewarded, and that while his inappropriate movements are sometimes followed by teacher reprimands, they are more often ignored or rewarded by peer approval or positive interaction.

The data can be incorporated into the body of the report (as in the preceding example) or described in tabular form in an Appendix to the report. Regardless of how the data are presented, their inclusion strengthens the validity of the conclusions drawn from the observation.

Table 1
Major Questions Guiding the School Observation Report

Descriptive information
What is the child's name, date of birth, parents' names, and contact information?
Who is the referring clinician?
What is the location, date and time of the actual observation?
What is the name of the teacher or any other staff involved with the child?
Reason for observation
Who requested the observation?
What are the goals of the observation?
What are the primary referral questions?
What are the specific target behaviors (both adaptive and maladaptive)?
Were any previous observations performed? What were the results?
Teacher interview
What problem behaviors are reported by teacher?
What academic difficulties are reported by teacher?
In what settings do the problem behaviors most frequently occur?
What are the suspected triggers of the problem behaviors?
What are the current consequences of the problem behaviors?
What past or current interventions were implemented by the teacher or other professionals? How effective were they?
Classroom environment
What is the number of students and staff in the classroom?
What is the size and shape of the classroom?
What is the location of furniture/equipment, seating arrangements and placement of the child?
Are there distracting stimuli, background noises, or outside class interruptions?
Are the traffic patterns well-defined and safe?
Are there established routines for toileting, drinks, snack time, etc.?
Are the class rules and consequences posted in a visible location? Are the rules reinforced?
Child observation
What classes, lessons, or tasks occurred during the observation period?
What evaluation procedures were used and are they described?
What were the results of the observation?
Describe the severity, frequency, and duration of the target behaviors
Describe the antecedents and consequences of the observed behaviors
Support all statements with data from the observation
Are the results of the observation reliable and valid?
Was the child's behavior during the observation representative of this child's behavior in this context more generally?
Recommendations
Based on the results, what should the teachers and school personnel do to effectively modify the child's behavior?
Based on the results, what can the parents do to effectively modify the child's behavior?
What implications do the results have for the therapist treating the child?
Is a follow-up school observation warranted? If so, when?

Using the Data to Make Recommendations

The primary purpose of the school observation is to provide data that will be useful to clinicians, teachers, and parents in understanding and modifying child behaviors. The mechanism through which data from the observation are put to actual use is the recommendations that appear at the end of the written report. There should be a natural progression from the primary referral questions, to the observed data, to the main conclusions, to the recommendations. The recommendations typically should not deviate from the primary referral questions and should not go beyond the data collected. For instance, consider the following recommendations, one not supported by data and thus making nonspecific recommendations:

> Regarding Alex's classroom behavior, it is recommended that his teacher or therapist work with him to help modify his current behavior problems.

and one based on the obtained data and making specific recommendations:

> Given Alex's out-of-seat behavior appears to be maintained by a lack of positive consequences for in-seat behavior, by the presence of positive peer interactions while out of his seat, and by the lack of negative consequences for being out of his seat, it is recommended that: (1) Alex's teacher use specific, labeled praise and, if needed, a contingency management program for increasing in-seat behavior, (2) specific commands and, if needed, time-out procedures for noncompliance, and (3) the teacher consider the use of a classroom-wide contingency program or consequence-sharing program to encourage in-seat behavior and discourage positive feedback for out-of-seat behavior.

Of course, more detail would be provided in the actual recommendations for each example. However, it is clear that congruence between the data and the recommendations offered and the incorporation of evidence-based techniques enhance the utility of the school observation and resulting report.

Follow-Up Observations

After the initial school observations, follow-up observations may be warranted for the purposes of providing an objective evaluation of the effectiveness of any intervention efforts and for identifying any additional behavior problems that may be present. Using a new observer will decrease the risk of reactivity and will increase the objectivity of the ratings by controlling for biases associated with the initial observer. However, using a new observer may introduce additional variance to the observation procedure, thus decreasing the reliability of ratings. That is, it could be that a given target behavior did not change in frequency from an initial observation to a follow-up observation; however, it appears to have changed due to differences in the accuracy of two different observers. Nevertheless, the use of follow-up observations often is warranted and is a useful tool for providing ongoing intervention to the child.

Conclusion

The school setting provides a naturally structured and contained environment with opportunities to assess children in multiple domains of functioning. Despite the prevalence of child behavior problems and the potential usefulness of conducting school observations, few practitioners have received training, or are proficient, in the use of these techniques. The purpose of this paper is to provide clinicians with an understanding of the theoretical rationale and a guide to assist in the practical implementation of school observations.

Several coding systems for the direct observation and assessment of school behavior have been developed and evaluated. Although of great importance for research and clinical efforts, these systems share several clinical limitations. Perhaps most importantly, they do not allow for the flexibility often needed in clinical settings. This article describes methods for conducting an idiographic direct observation of school behavior that is useful to clinicians conducting the observations, as well as to those who work with children or families with problems in the school setting. There is a trade-off inherent in the attempt to use an idiographic, tailored approach over one that is more uniform, in that it is much more difficult to assess the use of the former in a controlled manner (e.g., Abikoff, 2001). Thus, while the methods outlined in this article will be useful to practitioners working with children and families, we acknowledge their inferiority to existing coding systems regarding the presence of research evidence on their behalf. On balance, the methods outlined above are all based on techniques and procedures with decades of support in the behavior therapy literature (e.g., functional assessment, direct observation of classroom behavior).

This paper is intended to fill a gap in the clinical literature between the highly structured, research-based procedures commonly used to evaluate classroom behavior and the less structured, more subjective techniques often used to assess child behavior in such settings. We hope this paper provides clinicians with the tools necessary to perform scientifically informed, reliable, and valid evaluations of child behavior in school settings, and inspires addition clinical and research efforts in this important area.

References

Abikoff, H. (2001). Tailored psychosocial treatments for ADHD: The search for a good fit. *Journal of Clinical Child Psychology, 30*, 122–125.

Abikoff, H., Courtney, M., Pelham, W. E., & Koplewicz, H. S. (1993). Teachers' ratings of disruptive behaviors: The influence of halo effects. *Journal of Abnormal Child Psychology, 17*, 519–533.

Abikoff, H., & Gittelman, R. (1985). Classroom observation code: A modification of the Stony Brook code. *Psychopharmacology Bulletin, 21*, 901–909.

Abikoff, H., Gittelman-Klein, R., & Klein, D. F. (1977). Validation of a classroom observation code for hyperactive children. *Journal of Consulting and Clinical Psychology, 45*, 772–783.

Abikoff, H., Gittelman, R., & Klein, D. F. (1980). Classroom observation code for hyperactive children: A replication of validity. *Journal of Consulting and Clinical Psychology, 48*, 555–565.

Achenbach, T. M. (1986). *Manual for the Child Behavior Checklist/4-18 & 1991 Profile*. Burlington, VT: Department of Psychiatry, University of Vermont.

Achenbach, T. M., McConaughy, S. H., & Howell, C. T. (1987). Child/adolescent behavioral and emotional problems: Implications of cross-informant correlations for situational specificity. *Psychological Bulletin, 101*, 213–232.

Barkley, R. A., DuPaul, G. J., & McMurray, M. B. (1990). A comprehensive evaluation of attention deficit disorder with and without hyperactivity. *Journal of Consulting and Clinical Psychology, 58*, 775–789.

Blader, J., Nissen, S., Fleiss, K., & Kurtz, S. M. S. (2000). *Functional behavioral assessments and intervention plans: Program handbook*. New York: New York University Child Study Center.

Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multi-trait multi-method matrix. *Psychological Bulletin, 56*, 81–105.

Cohen, P., & Kasen, S. (1999). The context of assessment: Culture, race, and socioeconomic status as influences on the assessment of children. In D. Shaffer, C. P. Lucas, & J. E. Richters (Eds.), *Diagnostic assessment in child and adolescent psychopathology*. New York: The Guilford Press.

Cole, D. A. (1990). The relation of social and academic competence to depressive symptoms in childhood. *Journal of Abnormal Psychology, 99*, 422–429.

Durand, V. M., & Crimmins, D. B. (1988). Identifying the variables maintaining self-injurious behavior. *Journal of Autism and Developmental Disorders, 18*, 99–117.

Edelbrock, C. S. (1988). Informant reports. In E. S. Shapiro & T. R. Kratchwell (Eds.), *Behavioral assessment in schools: Conceptual foundations and practical applications* (pp. 351–383). New York: The Guilford Press.

Edelbrock, C. S., Costello, A. J., Dulcan, M. K., Conover, N. C., & Kalas, R. (1986). Parent-child agreement on child psychiatric symptoms assessed via structured interview. *Journal of Child Psychology and Psychiatry, 27*, 181–190.

Haaga, D. A. F. (2000). Introduction to the special section on stepped care models in psychotherapy. *Journal of Consulting and Clinical Psychology, 68*, 547–548.

Hanley, G. P., Iwata, B. A., & McCord, B. E. (2003). Functional analysis of problem behavior: A review. *Journal of Applied Behavior Analysis, 36*, 147–185.

Horn, W. F., Conners, C. K., Wells, K. C., & Shaw, D. (1986). Use of the Abikoff Classroom Observation Coding System on a children's inpatient psychiatric unit. *Journal of Psychopathology and Behavioral Assessment, 8*, 9–23.

Individuals with Disabilities Education Improvement Act of 2004, Pub. L. No. 108-446, 118 Stat. (2004).

Ivens, C., & Rehm, L. P. (1988). Assessment of childhood depression: Correspondence between reports by child, mother, and father. *Journal of the American Academy of Child and Adolescent Psychiatry, 27*, 738–741.

Jacobs, J. R., Boggs, S. R., Eyberg, S. M., Edwards, D., Durning, P., Querido, J. G., McNeil, C. B., & Funderburk, B. W. (2000). Psychometric properties and reference point data for the Revised Edition of the School Observation Coding System. *Behavior Therapy, 31*, 695–712.

Jensen, P. S., Watanabe, H. K., Richters, J. E., Roper, M., Hibbs, E. D., Salzberg, A. D., & Liu, S. (1996). Scales, diagnoses, and child psychopathology: II. Comparing the CBCL and the DISC against external validators. *Journal of Abnormal Child Psychology, 24*, 151–168.

Kandel, D. B., & Davies, M. (1986). Adult sequelae of adolescent depressive symptoms. *Archives of General Psychiatry, 43*, 255–262.

Kashani, J. H., Orvaschel, H., Burk, J. A., & Reid, J. C. (1985). Informant variance: The issue of parent-child disagreement. *Journal of the American Academy of Child and Adolescent Psychiatry, 24*, 437–441.

Kazdin, A. E. (1978). Artifact, bias, and complexity of assessment: The ABCs of reliability. *Journal of Applied Behavior Analysis, 10*, 141–150.

Kazdin, A. E. (2001). *Behavior modification in applied settings* (6th edition). Stamford, CT: Wadsworth.

Lambert, N. M., & Sandoval, J. (1980). The prevalence of learning disabilities in a sample of children considered hyperactive. *Journal of Abnormal Child Psychology, 8*, 33–50.

Ledingham, J. E., Younger, A., Schwartzman, A., & Bergeron, G. (1982). Agreement among teacher, peer, and self-ratings of children's aggression, withdrawal, and likeability. *Journal of Abnormal Child Psychology, 10*, 363–372.

Loeber, R. (1990). Development and risk factors of juvenile antisocial behavior and delinquency. *Clinical Psychology Review, 10*, 1–41.

McNeil, C. B., Eyberg, S. M., Eisenstadt, T. H., Newcomb, K., & Funderburk, B. (1991). Parent-child interaction therapy with behavior problem children: Generalization of treatment effects to the school setting. *Journal of Clinical Child Psychology, 20*, 140–151.

Miltenberger, R. G. (1997). *Behavior modification: Principles and procedures*. Pacific Grove, CA: Brooks/Cole.

Nock, M. K., Goldman, J. L., Wang, Y., & Albano, A. M. (2004). From science to practice: The flexible use of evidence-based treatment procedures in clinical settings. *Journal of the American Academy of Child and Adolescent Psychiatry, 43*, 777–780.

Nock, M. K., & Prinstein, M. J. (2004). A functional approach to the assessment of self-mutilative behavior. *Journal of Consulting and Clinical Psychology, 72*, 885–890.

Patterson, G. R., & Forgatch, M. S. (1995). Predicting future clinical adjustment from treatment outcome and process variables. *Psychological Assessment, 7*, 275–285.

Prinstein, M. J., Nock, M. K., Spirito, A., & Grapentine, W. L. (2001). Multi-method assessment of suicidality in adolescent psychiatric inpatients: Preliminary results. *Journal of the American Academy of Child and Adolescent Psychiatry, 40*, 1053–1061.

Reynolds, C. R., & Kamphaus, R. W. (1998). *Behavior Assessment System for Children: Manual*. Circle Pines, MN: American Guidance Service.

Shaffer, D., Fisher, P., Dulcan, M. K., Davies, M., Piacentini, J., Schwab-Stone, M. E., Lahey, B. B., Bourdon, K., Jensen, P. S., Bird, H. R., Canino, G., & Regier, D. A. (1996). The NIMH Diagnostic Interview Schedule for Children Version 2.3 (DISC-2.3): Description, acceptability, prevalence rates, and performance in the MECA Study. Methods for the Epidemiology of Child and Adolescent Mental Disorders Study. *Journal of the American Academy of Child and Adolescent Psychiatry, 35*, 865–877.

Shapiro, E. S., & Kratochwill, T. R. (Eds.). (2000). *Behavioral assessment in schools: Theory, research, and clinical foundations* (2nd ed.). New York: The Guilford Press.

Skinner, C. H., Dittmer, K. I., & Howell, L. A. (2000). Direct observation in school settings: Theoretical issues. In E. S. Shapiro & T. R. Kratochwill (Eds.), *Behavioral assessment in schools: Theory, research, and clinical foundations* (2nd ed.; pp. 19–46). New York: The Guilford Press.

U.S. Department of Health and Human Services. (1999). *Mental health: A report of the Surgeon General—Executive summary*. Rockville, MD: U.S. Department of Health and Human Services, Substance Abuse and Mental Health Services Administration, Center for Mental Health Services, National Institutes of Health, National Institute of Mental Health.

van Lier, P. A., Verhulst, F. C., & Crijnen, A. A. (2003). Screening for disruptive behavior syndromes in children: The application of latent class analyses and implications for prevention programs. *Journal of Consulting and Clinical Psychology, 71*, 353–363.

Address correspondence to Matthew K. Nock, Ph.D., Department of Psychology, Harvard University, 33 Kirkland Street 1280, Cambridge, MA 02138; e-mail: nock@wjh.harvard.edu.

Received: February 24, 2005
Accepted: March 17, 2005