2

Procedural Issues Associated with the Behavioral Assessment of Children

CHRISTOPHER H. SKINNER
JENNIFER T. FREELAND
EDWARD S. SHAPIRO

A variety of reasons may account for the increased interest among educational and psychological professionals in behavioral assessment that has occurred during the last quarter of the 20th century (Shapiro & Kratochwill, 2000). Recently, legal changes related to inappropriate classroom behaviors have resulted in an increase in the need for training in behavioral assessment theory and practice. Specifically, the 1997 Amendments to the Individuals with Disabilities Education Act (IDEA) require that individual education plan (IEP) teams address behavioral challenges of students with disabilities. These changes include requirements for conducting functional behavioral assessments for the purpose of intervention development. Furthermore, schools are charged with using postintervention data to evaluate intervention effectiveness and alter interventions as needed (Nelson, Roberts, Rutherford, Mathur, & Aaroe, 1999; Yell & Shriner, 1997). Both the language (e.g., functional assessment) and the emphasis on linking assessment to intervention is causing increased attention to behavioral assessment procedures.

The amendments to IDEA were preceded and influenced by advancements in behavioral assessment research and procedures. Specifically, methodological advances in behavioral assessment procedures (e.g., experimental functional analysis and descriptive functional assessment) have been shown to be effective for producing data that lead practitioners and researchers to effective interventions (Iwata, Dorsey, Sifler, & Richman, 1982; Lalli, Browder, Mace, & Brown, 1993; McComas, Hoch, & Mace, 2000; O'Neill et al., 1997).

Additional reasons for an increase interest in behavioral assessment procedures include a dissatisfaction with indirect measures, frustration by consumers with the failure of traditional models of assessment to lead to more effective interventions, a lack of accountability at the individual student level associated with the use of indirect measures, and a desire for more flexible assessment procedures that can serve a variety of purposes (Reschly & Ysseldyke, 1995; Shapiro, 1987; Telzrow, 1999). A broader more inclusive conceptual framework of behavioral assessment procedures that include more traditional assessment methodologies also may have increased the number of professionals, researchers, and educators inter-

ested in behavioral assessment (Cone, 1978; Nelson & Hayes, 1979, 1986).

Regardless of the reasons for the increased interest in behavioral assessment procedures, psychologists, educators, counselors, social workers, and others who work with children are likely to be expected to be able to understand and apply behavioral assessment procedures (Barnett et al., 1999; Hendrickson, Gable, Conroy, Fox, & Smith, 1999). In the next section, we briefly describe characteristics of behavioral assessment procedures that may serve to differentiate them from other more traditional assessment procedures. More traditional assessment procedures such as interviews, checklists, and rating scales have been incorporated into a behavioral framework and applied to the assessment and treatment of children (see Shapiro & Kratochwill, 2000). Because other chapters in this series cover many of these procedures, this chapter focuses on direct behavioral assessment procedures (i.e., direct observation in natural and analogue environments) and issues related to the application of these procedures across behaviors and settings.

ASSUMPTIONS REGARDING CAUSES OF BEHAVIOR

Hartmann, Roper, and Bradford (1979) provided a comprehensive overview that contrasts behavioral and traditional assessment with respect to assumptions, implications, use of data, and other characteristics. Each of the differences can be traced to primary assumptions associated with causes of behavior. Under a behavioral model, behaviors are caused and/or maintained by current environmental conditions and past learning history. However, with more traditional models of psychology, behavior is seen as caused by intrapsychic or within-child traits, conditions, or mediating variables.

Under more traditional models of assessments, overt behavior is measured and used to infer these within-child variables assumed to cause these behaviors. Thus, a child's inappropriate behavior may be caused by his passive–aggressiveness, attention-deficit disorder, or faulty cognitions. Because these conditions that are thought to cause behaviors are also seen as relatively stable, traditional models of assessment have a history of being used for identifying, classifying, or diagnosing problems and predicting future behavior (Hartmann et al., 1979).

Under a behavioral model, behavior is not seen as a mere symptom of some other underlying (i.e., within-child) problem. Rather, behaviors are viewed as legitimate problems in and of themselves. However, the most important distinction may be that behavioral theory does not make large leaps of influence and attribute the cause of a child's behavior to some underlying construct. Instead, current behaviors, which are directly observable are maintained by current environmental events (e.g., antecedent and consequent stimuli) that in many instances are also directly observable. Thus, under a behavioral model, both the behaviors of interest and the events thought to maintain behaviors can be assessed more directly, regardless of what procedures are used to measure the behavior and/or environmental events that may maintain the behaviors (Carr, 1993; Gresham, 1998; Skinner, Dittmer, & Howell, 2000).

In addition to lending themselves to direct assessment, environmental variables that are thought to maintain behaviors under a behavioral model are often mutable. For this reason, behaviorists have been less likely to focus all their assessment efforts on identifying and measuring problems or problem behaviors. Instead, behavioral psychologists are equally if not more concerned with developing procedures for identifying and measuring environmental events that are thought to cause these behaviors (Carr, 1993; Gresham, 1998). Once variables that cause problem behaviors are identified, interventions designed to alter target behaviors that are based on assessed causes of these behaviors can be developed (e.g., McComas et al., 2000; Myerson & Hale, 1984).

BEHAVIORAL ASSESSMENT PROCEDURES

Shapiro and Browder (1990) present a continuum of behavioral assessment procedures from most direct to least direct. Indirect as-

sessment procedures include self-report and informant report measures (e.g., checklists, rating scales, and interviews) that are covered in other chapters in the current series. Direct behavioral assessment procedures include direct observation in natural settings, direct observation in analogue settings, and self-monitoring. Additional direct observation procedures require peers, parents, teachers, or others who are a part of the child's natural environment to observe and record data (Skinner, Rhymer, & McDaniel, 2000). With each of these procedures, behaviors can be observed and recorded as they occur. Furthermore, each of these direct assessment procedures can employ similar data collection procedures. The primary difference across procedures is who collects the data (independent observers, parents, teachers, peers, or target children themselves) and under what conditions (natural vs. analogue or artificial environmental conditions). Next, we describe general procedures for directly observing and recording behaviors. Then, we describe issues related to specific direct observation data collection and procedures.

RECORDING DIRECT OBSERVATION DATA

A variety of procedures can be used to record direct observation data. These procedures have relative strengths and weakness related to (1) the goal of data collection; (2) the rates, topography, and duration of target behaviors; (3) the conditions under which data collection must occur; and (4) physical, temporal, and resource constraints related to data collection.

Narrative Recording Procedures

Benefits of Narrative Recording Procedures

The least structured form or narrative recording merely requires an observer to write narrative descriptions of behaviors and sometimes events surrounding those behaviors. Narrative recording procedures are often used to communicate general information about a child's behavior (e.g., Johnny seemed tired today). Narrative recordings can also be used in the initial stages of problem solving to help (1) identify, validate, or confirm target or problem behaviors; (2) to form a general idea of problem behaviors rates, intensity, and topography; and (3) to begin to identify variables that may be serving to maintain target or problem behaviors (Skinner, Rhymer, & McDaniel, 2000). Although there are numerous narrative recording procedures and a variety of reasons to collect narrative recordings, daily communication logs, descriptive time sampling, and Antecedent–Behavior–Consequence (A–B–C) analysis provide a fairly broad overview of these procedures.

Daily communication logs can allow one to record data that can be used to track a child's general progress or to communicate with others who also work with a child. For example, a teacher working with students with autism may write a general description of a child's school behavior to be sent home to the child's parents. A parent of a child with social–emotional problems may be asked to make similar recordings for a child's therapist (e.g., Johnny appears to be becoming less anxious). These narrative descriptions are useful because individuals can provide flexible and rich descriptions of a child's behavior that may assist others who are working with that child.

If specific behaviors are of concern (e.g., self-injurious behavior during school), narrative recordings can provide data about the general topography (e.g., scratching self vs. slapping self), intensity (hitting self hard), rate (hitting self infrequently), or duration of a behavior. These data can prove useful when developing more systematic observation procedures. For example, descriptions of the topography of a behavior can be used to help develop operational definitions. Data on behavior rates and durations are useful when one is developing interval recording procedures. Furthermore, narrative recordings can provide a general indication of behavior variability and the conditions when the target behavior is likely to occur (hitting self only in the first hour of school). These data can be used to determine when observers may be more likely to have the opportunity to directly observe target behaviors. In addition, because data on variability may provide some initial information regarding conditions that may be serving to maintain target behaviors (e.g., is

disruptive during independent seat-work time), these data could prove useful for developing intervention procedures.

Narrative recordings are often used to keep records of specific low-rate behaviors or events. Following unusual or dangerous events, teachers, parents, or staff often use narrative recording when completing "incident reports." When behaviors are clear and obvious and occur at low rates, incident reports can be converted to frequency counts of specific behaviors or events (e.g., number of times a child became physically aggressive). In this manner, narrative recordings can sometimes be used to determine whether students' behaviors are increasing, decreasing, or remaining stable.

Although narrative recording can sometimes be translated into empirical data for low-rate behaviors or events, narrative time sampling may allow one to collect data on more typical behaviors and conditions. Narrative time sampling requires observers to write a narrative recording of a student's behavior at predetermined intervals. For example, every 2 minutes an observer could look at a child and then write what the child was doing at the time and record other environmental conditions that were present at that time. This type of momentary time sampling can provide an estimate of rates of specific behaviors (Shapiro & Skinner, 1990).

A–B–C narrative recording requires observers to record narrative descriptions of both behaviors of interest and antecedent and consequent conditions that may be functionally related to those target behaviors. Typically, observers use the occurrence of target behaviors as a cue to write a description of the (1) target behavior; (2) general antecedent conditions (e.g., class completing independent seat-work assignment involving punctuation) and specific antecedent events (e.g., a peer takes a students pencil), that precede target behaviors; and (3) consequent events or events that immediately followed the target behavior.

Limitations of Narrative Recording Procedures

Although narrative recording procedures can be extremely useful for identifying or verifying behaviors and environmental variables that may be maintaining those behaviors, there are several limitations associated with narrative recording procedures. By definition, narrative recordings involve nonsystematic data recording procedures. Although this method of recording may allow for rich, flexible data collection, narrative recordings often yield imprecise data. For example, an observer may spend some time with a family and write that Johnny disobeyed his parents often. Without a more precise definitions of the words "disobey" and "often" it is difficult to use these data to determine what was actually occurring during the observation period. In addition, the quality and quantity of narrative data are likely to be influenced by many variables, including the observer's writing skills (e.g., writing speed and vocabulary) and his or her training in recognizing and recording environmental events that may be related to target behaviors (e.g., recognizing that shifting from one type of mathematics problem to another may occasion escape–avoidance behaviors). Unless behaviors or events of interest are extremely obtrusive or obvious, discreet, uncomplicated, and occur at low rates, it is often difficult, if not impossible, to verify data collected via narrative recordings (Shapiro & Skinner, 1990).

A second major limitation of narrative recording procedures is related to the time required to write narrative descriptions. Because it takes considerable time to write narrative descriptions, observers cannot record all behaviors or events observed. Furthermore, because it takes time and attention to write narrative description, narrative recording procedures often require discontinuous observation (i.e., observers halt observation in order to record data). Discontinuous observation has several limitations. Because the observer is not observing behaviors continuously (i.e., they pause to write out their narratives) observers may miss the opportunity to observe and record important antecedent or consequent events that could lead to effective interventions (Skinner, Dittmer, & Howell, 2000). For example, while busy recording target behaviors, observers may fail to observe an event that occurred after the target behavior and is reinforcing that target behavior. Fortunately, a variety of empirical recording procedures have been developed that can ad-

dress many of the limitations associated with narrative recordings.

Empirical Recording

Empirical recording procedures yield more precise measures of behavior than narrative recoding procedures. This precision can allow for the fine distinction necessary for measuring the effects of interventions over brief periods. In addition, empirical recording procedures yield data that can be independently verified. Verification is especially useful when (1) important decisions are being made (e.g., is the child's self-injurious behavior getting worse or is it severe enough to warrant alternative placement), (2) primary observers may be susceptible to biases (e.g., the residential staff member who has recently had a physical altercation with the student is collecting data), or (3) when the behaviors being measured are difficult to observe and record.

Empirical recording systems typically require observers to record tally marks in appropriate columns or interval blocks when target behaviors are observed. Because this data recording procedure is much more time efficient than writing a narrative description, empirical recording often allows observers to collect data in a more continuous manner and also record observations across a greater array of behaviors or events in a relatively brief period of time. These advantages may make it easier to precisely measure target behaviors. In addition, efficient recording may increase the probability of observers being able to observe and record important antecedent and consequent events that may be maintaining these behaviors.

Operational Definitions

To record the presence of a target behavior, the behavior must first be operationally defined. Operational definitions are typically based on the topography or shape of the behavior. In some instances the intensity or duration of behaviors is also included in operational definitions. For example, "inappropriate voice" may include speaking too loud or too soft (intensity) and "passive–aggressive behavior" may include complying with mothers directions within 5 seconds of the direction being issued (duration). However, to avoid imprecise data collection, operational definitions rarely include inferred characteristics of the behavior such as intent or purpose of the behavior (Skinner, Dittmer, & Howell, 2000).

Several procedures are useful for developing operational definitions of target behaviors. The typical procedure is to collect descriptive data from the referring agent (e.g., teacher or parent). Often referrals are accompanied by broad, vague descriptions of problem behaviors and interviews can be used to form more precise definitions of target behaviors (Bergan & Kratochwill, 1990). Including examples of behaviors that fit an operationally defined behavior and those that do not fit can help clarify operational definitions (see Saudargus, 1992). In other instances, it is easier to collect direct observation data when target behaviors are operationally defined to include an entire class of behaviors. For example, physical aggression against peers could include any instance of biting, hitting, or kicking a peer. However, it is difficult and often not meaningful to record each particular bite, hit, or kick. Instead, aggressive *instances* could be defined as beginning when a child engages in any of the previously mentioned behaviors and ending with the absence of those behaviors over a 2-minute interval.

When target behaviors are unusual or idiosyncratic (e.g., hand flapping), using narrative recording procedures to develop a description of the behavior may help one develop operational definitions. This observation time may also help one obtain a clearer picture of the duration, rate, and continuity of the target behaviors which is often necessary when developing direct observation recording procedures. Finally, it is not always necessary to construct a new operational definition for every referred problem. In many instances, it may be possible to find appropriate operational definitions in empirical intervention journal articles, behavioral psychology text, and structured codes.

Data Recording Procedures

Developing direct observation systems also requires the development of data recording procedures. Successfully matching the data

recording procedures to the target behavior can make data collection more reliable and enhance the social, educational, and clinical validity of the data collected.

Event Recording

Event recording requires an observer to record the number of times a behavior occurs during an observation session. Event recording is often used to collect data on discrete behaviors that have a clear beginning and end (e.g., leaving one's seat). However, because every instance of the behavior is recorded, it can be difficult to collect data on behaviors that occur at high rates (e.g., pencil tapping).

Event data can be reported as frequency counts (e.g., Johnny did not comply with his father's directions on 20 occasions and Jane had six aggressive outburst). However, converting data to either rate or percentage data often provides more useful information. For example knowing that Johnny failed to comply with his father's directions 10% of the time or 20 out of 200 requests provides much more meaningful data regarding Johnny's compliance. Reporting frequency counts as rate data (e.g., six aggressive outbursts for the schoolweek vs. a schoolday) makes frequency counts more meaningful. Converting frequency counts to rate or percentage data allows one to compare data across observation periods when opportunities to engage in behaviors or interval lengths are unequal. Thus, these types of conversions are extremely useful when repeated measures data are used to (1) analyze behavior trends, (2) analyze behavior variability, and/or (3) compare a child's behavior across conditions (e.g., baseline vs. treatment phases).

Duration Recording

Whereas event recording is often used for discrete behaviors that occur at moderate or low rates, duration recording can be used to collect data on continuous behaviors or behaviors that occur at extremely high rates. It is possible for an observer to use a stopwatch to record the amount of time a child spends engaged in specific behaviors. However, it is often difficult to start and stop a stopwatch in a reliable manner. Furthermore, with continuous behaviors it can be difficult to reliably determine when a behavior begins and ends. For example, if on-task were defined as head oriented toward assignments or a speaker, an observer would be constantly starting and stopping the stop watch every time the student reoriented.

Fortunately, time-sampling procedures allow observers to collect reliable duration estimates. There are three types of time sampling (i.e., momentary, whole interval, and partial interval) and each requires observers to record behaviors on an interval-by-interval basis. Although observers can use a clock, stopwatch, or wristwatch to mark intervals, using audiotapes to mark intervals may make it easier to observe and record student behavior. Marking intervals with audiotapes may also make it easier to synchronize recording intervals when direct observation data are being verified by another observer (e.g., when collecting interobserver agreement data).

Momentary time sampling requires observers to record the presence of a target behavior if it is occurring at the moment the interval begins. When whole interval time sampling is used, observers record the presence of a target behavior only if it occurs for the entire interval. With partial interval time sampling, an interval is scored if the behavior occurs at any point during an interval.

Momentary time sampling is often used when data collection systems are complex and require observers to record a variety of behaviors and events. Whole and partial interval time sampling often require more continuous observation (e.g., observe for entire intervals). Thus, these procedures are less useful when observers are attempting to collect data on multiple behaviors and events or across multiple people (e.g., several peers of a child and the child's parents and siblings).

Time-sampling methods involve sampling and therefore are susceptible to sampling error. Whole interval time sampling tends to underestimate behavior durations and partial interval time-sampling procedures overestimate behavior durations (Lentz, 1982; Powell, Martindale, Kulp, Martindale, & Bauman, 1977). The length of observation intervals can also affect time-sampling estimates. Longer intervals tend to amplify overestimates of partial interval time sam-

pling and shorter intervals will tend to underestimate estimates provided by whole interval time sampling. When momentary time sampling is being used, shorter intervals would increase the size of the sample collected during an observations session, thus reducing sampling error of the duration estimates.

Because time sampling provides estimates of durations, when reporting data observers should not indicate the percentage of time the students was engaged in target behaviors. For example, it would be inaccurate to report that Jim was engaged in self-injurious behavior 80% of the time. Instead one should report the length of the observation sessions (e.g., 10 minutes), the length of the intervals (e.g., 20 seconds), the type of time sampling recording being used (e.g., momentary time sampling), the number of intervals the behavior was recorded, and the number of intervals the behavior was not recorded (e.g., self-injurious behavior was recorded for 24 of 30 intervals) and the percentage of intervals the behavior was recorded (e.g., self-injurious behavior was recorded for 80% of the observed intervals).

Interval Recording, Sequential Events, and Direct Observation Systems

Interval recording can be used to estimate the duration of time children are engaged in specific behaviors. Frequency data can also be recorded using intervals. This is especially useful when observers are collecting data on both target behaviors and variables thought to be maintaining those behaviors in natural environments. Recording events that occur during intervals provide a more precise record of the sequence of events. Thus, using interval recording observers can identify events that occurred immediately prior to and following target behaviors (see Saudargus, 1992). To the extent that these events are serving to occasion and maintain target behaviors, this type of recording may provide data that lead to more effective interventions.

OBSERVERS AND ENVIRONMENTS

Several issues are related to direct observation that can affect the quality and clinical utility of direct observation data. These issues are affected by who collects the data (e.g., independent observer, participant observer, peer observer, and/or self-observer) and under what conditions data are collected (i.e., natural environments vs. analogue environments).

Independent Observer Collecting Data in Naturalistic Settings

Having an independent observer enter a child's natural setting (e.g., home or school) and record the child's behavior and environmental events surrounding those behaviors is the most direct and often considered the most desirable form of behavioral assessment (Cone, 1978; Hintze & Shapiro, 1995). However, there are several major concerns about using independent observers to collect data in a child's natural environments. Perhaps the most important concern is that the process of the data collected will affect the child's and others' (e.g., parent, teacher, and peers) behavior. This process is known as reactivity. A variety of variables including (1) obtrusiveness of the observer, (2) perceived power or role of observer, and (3) what the child is told about the observer's presence may affect reactivity (Johnson & Bolstad, 1973). However, it is not possible to predict or measure the impact that reactivity is having on behavior(s) during any particular observation session (Shapiro & Skinner, 1990). Therefore, when the goal is to collect assessment data that most accurately reflect naturally occurring conditions, observers should make efforts to reduce reactivity.

First, efforts should be made to reduce the conspicuousness of the observer. One-way mirrors are an excellent device that allow one to collect data in an inconspicious manner. However, often independent observers cannot conceal their presence in the child's natural environment. When this is the case, children are likely to ask questions about the observer. In these instances, teachers, parents, or others who are supervising the children should not lie to the children, but they should avoid providing too much information. For example, informing a class of fourth-grade students that the observer is there to collect data on John because people are concerned about his inappropriate be-

havior could be both unethical and illegal but clearly is likely to cause reactivity. Not only is John's behavior likely to be affected, John's peers are likely to interact differently with John. Thus, it may be best to provide vague, general statements about why the observer is present. For example, children could be told that Mrs. Smith is here to observe and learn more about fourth-grade classrooms.

Children often attempt to interact with independent observers (Johnson & Bolstad, 1973). Observers should not respond to students, as even quick social interactions may encourage children to continue initiating other interactions. These interactions with observers who would not be present if they were not collecting direct observation data are obvious examples of reactivity. Not all interactions are verbal. Children may engage in nonverbal behavior (e.g., writing on a piece of furniture) while they know that an independent observer is watching them in order to see if the observer reacts (e.g., tell their parents and take away their pen). If observers react, the child's future behavior is likely to be strongly influenced by the presence or absence of the observer. Of course, there are exceptions to this rule. If there is a clear and present danger (e.g., a child is leaning far out an open window or the observer sees a child pointing a knife at another child) observers are legally and ethically obligated to react.

Several other procedures can be used to reduce reactivity (Johnson & Bolstad, 1973; Kazdin, 1977; Skinner, Dittmer, & Howell, 2000). Observers should not orient themselves directly toward the target child or children. Staring directly at target children can only increase reactivity. Observers should sit in an area of the room where they are less likely to be noticed (e.g., the back of the classroom). When possible, observers should attempt to enter the natural environment at unobtrusive times (e.g., enter a classroom before students do). To children, it can appear that observers are not busy. This may increase the probability of children approaching the observer. Therefore, observers who appear extremely busy (e.g., constantly writing or quietly shuffling materials) may be less likely to occasion reactivity during observation periods.

Timing related to other assessment activities may also be related to reactivity. For example, psychologists often interview children as part of the assessment process. When the person who interviewed the specific child enters the classroom, that child is likely to behave differently (i.e., reactivity) because of the interview. In these instances, reactivity may be reduced by delaying the child interview until direct observation data are collected. Another solution would be to have one person conduct the interview and another collect the direct observation data.

Although independent observers may occasion reactivity, this reactivity is likely to subside over time as children and others in the environment return to their typical behavior. Thus, initial observation sessions may yield less naturalistic data than subsequent sessions. Video cameras can also be used to record data that observers can record at a later date. As with the presence of an outside observer, the presence of a video camera is likely to occasion high levels of initial reactivity (e.g., mugging for the camera and "Hi Mom") that tend to subside as children become acclimated to the equipment.

There are several advantages to using independent observers to collect direct observation data. Because external observers do not have a history of interacting with the child, they may be less susceptible to observer biases that could cause inaccurate or inconsistent data recording. Learning to collect direct observation data in a reliable manner can be time-consuming. It may be most efficient to train a few external observers and use them as direct observation specialists to collect data across environments (e.g., across classroom or homes). One way to make this system even more efficient would be to train these observers to collect data using preestablished direct observation systems (for examples of classroom observation systems, see Alessi & Kaye, 1983; Saudargus, 1992; Shapiro, 1996). These systems have been developed to allow observers to collect data on multiple behaviors, events, and conditions that may provide useful information for making decisions across problems and concerns. After becoming fluent with collecting data using these codes, observers can then make alterations in the system on a case-by-case basis (e.g., add a specific operational defini-

tion and recording procedure for a child's idiosyncratic behavior).

Teacher, Parent, or Other Internal Observer in Natural Settings

When target behaviors occur at low rates and/or unpredictable times it may not be practical for external observers to record direct observation data. For example, a child may have a tantrum about once every week where he throws and damages objects in his environment. Because this behavior occurs infrequently, it is difficult for an independent observer to schedule times when they can observe this behavior. In these instances, parents, teachers, or other adults who are part of the child's natural environment can be used to record direct observation data.

When working with teachers, parents, or others who are working or supervising children in their natural environments, it may be important to construct relatively simple recording systems that require little time to record data. For example, a teacher could keep a daily frequency count of a student's compliant and noncompliant behaviors by moving a penny from one pocket to another every time a student complied and moving a dime every time a student did not comply.

When interval recording systems are being used, manipulating interval length can also make data collection easier. For example, it is much easier for a parent to record whether a child wet the bed during an entire night than it is to record whether a child wet the bed during a specific hour interval. Altering operational definitions can also make collecting data easier. For example, rather than measuring a student's in-seat behavior by determining if one or both buttocks are in contact with the chair seat, one could record whether or not any part of the student's body is in contact with this desk or chair. In addition, care should be taken with procedures used to mark intervals. Using natural occurring intervals may reduce disruptions (e.g., intervals that last from waking to lunch, lunch to dinner, and dinner to bedtime or intervals that last an entire class period during schooldays). Furthermore, some procedures for marking intervals are likely to be less obtrusive and disruptive. For example, using a wristwatch beep as opposed to a cooking timer may allow a teacher to mark intervals in a manner that is less likely to disrupt typical classroom behaviors and routines.

When using parents, teachers, or other adults who are part of a child's natural environment to record observations, the process of data collection could have a great impact on how they interact with the target child. For example, suppose a parent is collecting data on a child's compliant and noncompliant behavior by moving pennies and dimes from one pocket to another. While collecting these data, the parent may come to realize that the child is rarely complying and reduce the number of requests or demands made on the child. This reduction in the number of demands may increase or decrease the child's compliance rate. Thus, the reactivity occasioned by parents or other internal observers collecting data can alter the child's natural environment and reduce the clinical utility of the data collection procedures.

Internal observers who have a history of interacting with a child may also be more susceptible to observer drift. For example, a parent who has listened to her child scream all day may be more (frustration) or less (habituation) likely to score a child's request as inappropriate verbal behaviors. Finally, parents, teachers, or others who work with a child may have a stake in the decisions being made based on that data. When this is the case, people who are making the decisions may be less likely to trust that data (e.g., a judge making a custody decision) when it is collected by a stakeholder (e.g., the father).

Peer Observation and Recording

Children can be used to collect direct observation data on the behavior of classmates, siblings, or peers. The reactivity associated with peer monitoring may occasion both desired and undesired behaviors, dependent on the behavior(s) being monitored. Within educational environments, tutoring programs often require peers to observe and record classmates academic behaviors (e.g., Greenwood et al., 1987). Peer monitoring is an efficient procedure that can allow for immediate evaluation of academic responses. This immediate evaluation can cue immediate error correction procedures that prevent

students from practicing inaccurate responding. In addition, immediate evaluation can serve as immediate reinforcement for accurate responding (McLaughlin & Skinner, 1996). Besides enhancing tutee's academic performance, the process of observing, recording, and evaluating peers' responses has been shown to improve the tutor's or monitor's academic skills (e.g., Skinner, Shapiro, Turco, Cole, & Brown, 1992). Although there may be a tendency to have older, more mature students, or students with stronger academic skills, serve as monitors or tutors, researchers have trained younger students, same-age peers, and students with disabilities to observe, record, and evaluate their peers' academic responses (Carden-Smith & Fowler, 1984; McCurdy & Shapiro, 1992). Serving as a peer monitor or tutor may also enhance a student's self-image and academic esteem (Henington & Skinner, 1998; Stern, Fowler, & Kohler, 1988).

Children can also observe, record, and report peers' incidental social behaviors. In many environments, incidental antisocial behaviors are punished. In some instances, children learn to avoid punishment by not performing those behaviors. In other instances, children learn to avoid being punished by avoiding being caught or observed performing these behaviors. For example, it would be unusual for some children to forcefully take a toy from a peer when a parent or teacher was obviously watching. Thus, in some instances peers may be the only people who observe a child's inappropriate behaviors. When these behaviors are serious (e.g., bringing a gun to school or torturing an animal), it is essential that children who observe these behaviors also report these behaviors.

Often children learn to monitor and report their peers' incidental antisocial behavior (i.e., tattle) without any specific or programmed instruction (Skinner, Cashwell, & Skinner, 2000). Therefore, it should be relatively simple to train and encourage children to observe and report peers' inappropriate behaviors. All methods of recording data described earlier could be used by students to record their peers' behavior. However, peers may have difficulty recording behaviors that occur at high rates or accurately recording behaviors using complex observation systems. In addition, having children write narrative reports can be time-consuming and the quality of these data may be compromised depending on students' writing (e.g., speed and vocabulary) and observation skills. Thus, in some instances, it may be best to train children to report peers incidental inappropriate behavior (e.g., tattle) and have adults record these behaviors (Henington & Skinner, 1998). One procedure used by an elementary school teacher was to have children describe their peers' behavior into a tape recorder.

There are several limitations associated with using peers to collect data on antisocial behaviors (Henington & Skinner, 1998; Skinner, Cashwell, & Skinner, 2000). Observing peer behavior can distract children's attention from their own behavior or their teacher's, parent's, or supervisor's instructions or directions. Peers may intentionally or unintentionally provide inaccurate reports of peers' behavior. Intentionally inaccurate reports can be used as a form of aggression to get other children in trouble. Such reports can also be used to shift the blame for a behavior to another peer. Peers may threaten their classmates to prevent them from reporting behaviors or physically assault them in retaliation for reporting those behaviors.

Requiring children to monitor inappropriate behaviors may teach children to focus on their peers' inappropriate behaviors to the exclusion of their peers' appropriate behaviors. This may reduce children's awareness of their peers' incidental prosocial behaviors and cause children to form negative impressions of their peers. Finally, it would be inappropriate to have peers collect data on the inappropriate behavior of a particular child, as this process would encourage peers to view this child as being deviant or bad and may increase the probability of children socially rejecting their peer (Cashwell, Skinner, Dunn, & Lewis, 1998; Ervin, Miller, & Friman, 1996).

Although there are some serious negative side effects associated with using children to collect data on peer's incidental inappropriate behaviors, having children observe, record, and report peers incidental prosocial behaviors could have several positive side effects. Researchers have attempted to use peer observation and peer reporting to im-

prove the social interactions and social status of socially rejected children and/or children who displayed high rates of inappropriate behaviors. Results have shown that having groups of children publicly report peers' incidental prosocial behaviors at daily group meetings can improve the student's social status, increase interactions, and increase cooperative interactions among students (Bowers, McGinnis, Ervin, & Friman, 1999; Ervin, Johnston, & Friman, 1998; Ervin et al., 1996; Jones, Young, & Friman, 2000; Robinson, 1998).

Although there are positive side effects associated with having children observe and report peers' incidental prosocial and academic behaviors, these side effects represent reactivity. It is not possible to separate behavioral changes caused by reactivity versus behavioral changes caused by planned (e.g., interventions) or unplanned changes in the student's immediate or temporally distant environments. Thus, even when peer monitors collect accurate and reliable data, interpretation of these data must be tempered with the possibility that these data do not reflect natural environmental conditions. Furthermore, the process of peer monitoring can have a significant impact on the behavior of target children and peer monitors. This impact can be significant enough to cause broad changes within the child's natural environment that may alter how others (teachers and parents) interact with the child. Because this reactivity may be unpredictable, the use of peer monitoring requires careful monitoring for reactivity across individuals in the environment and all decisions based on these data must be tempered with the understanding that reactivity may have altered naturally occurring behaviors.

Self-Monitoring

Self-monitoring requires children to observe and record their own behavior. There are several advantages to using self-monitoring over other direct observation procedures (Cole, Marder, & McCann, 2000). Self-monitoring is an efficient procedure for collecting data when other data collectors are unavailable or too busy. In addition, self-monitoring may be the only procedure that allows one to collect data that are difficult for others to observe, such as (1) cognitive or emotional behaviors (e.g., fear or cognitive steps taken to solve a mathematics problems), (2) behaviors that occur at low rates, (3) behaviors that occur at unpredictable times, and (4) behaviors that occur in settings that do not lend themselves to direct observation by others.

All previously described procedures for observing and recording data can be employed when self-monitoring is used. However, as with peer monitoring, several variables should be taken into account when structuring self-monitoring procedures. With young children and children with writing skills deficits, narrative recording may prove cumbersome and yield data that are difficult to interpret. It may also be difficult for children to observe and record several behaviors or events simultaneously (Nelson, 1977). Because children may find it time-consuming and difficult to consistently make subtle distinctions in behaviors, operational definitions should also be both clear and described in simple terms. In some instances, self-monitoring procedures can disrupt children who are engaged in desirable behaviors. For example, if a student is having problems maintaining her on-task behavior, interrupting that student at random intervals and requiring her to record whether she was reading silently could make it more difficult for this student to remain on task and finish assignments (Skinner & Smith, 1992).

As with all other direct observation procedures, reactivity and accuracy are a concern when children self-monitor. The quantity and quality of training can have an impact on the accuracy of self-monitoring data (e.g., Shapiro, McGonigle, & Ollendick, 1980). When children are told that someone else will also be monitoring and recording their behavior (e.g., checking their work or collecting interobserver agreement data), they may be more likely to self-record accurately (Santogrossi, 1974). Reinforcement for accurate self-monitoring also enhances accuracy (Fixsen, Phillips, & Wolf, 1972; Lloyd & Hilliard, 1989). The valence of the behavior being monitored may also enhance self-monitoring accuracy. Researchers have shown that children may be more likely to accurately observe and record their own appropriate, as opposed to inappropriate, behaviors (Nelson, Hay, De-

vany, & Koslow-Green, 1980; Nelson, Hay, Hay, & Carsten, 1977).

The process of observing and recording one's own behavior often causes reactivity. Fortunately, the direction of the changes bought about via self-monitoring are generally in the desired direction. Thus, when children are trained to observe and record inappropriate behaviors, those behaviors tend to decrease. When they are trained to observe and record their appropriate behaviors, those behaviors tend to increase. A variety of other variables can affect the degree of reactivity that occurs when children observe and record their own behavior, including (1) training in self-monitoring, (2) obtrusiveness of self-recording devices, (3) timing of self-monitoring, and (4) accuracy of self-recording (Glynn & Thomas, 1974; Nelson et al., 1980).

The reactivity caused by self-monitoring tends to make this a popular intervention. Not only is it efficient but self-monitoring is seen as moving children away from external control and encouraging students to become more actively involved in programs designed to maintain or alter their own behaviors (Kern, Marder, Boyajian, Elliot, & McElhattan, 1997). However, as with all forms of reactivity, it is difficult to predict the degree of behavior change that occurs in reaction to self-monitoring. For example, some studies have shown that accurate self-recording results in higher levels of reactivity (Shapiro, 1984). Other studies have shown that accurate recording is not necessary for reactivity to occur. In fact, some students who where trained to self-record but failed to ever record data showed changes in their behavior (Hayes & Cavoir, 1977). Because the presence or absence of reactivity cannot be reliably predicted or measured, assessment data collected via self-monitoring must always be interpreted with an understanding that these data may be affected by reactivity (Shapiro & Skinner, 1990).

Direct Observation in Analogue Settings

For a variety of reasons it is sometimes difficult to assess specific behaviors in natural environments. For example, some specific social skills or social behaviors occur at low rates or unpredictable times (e.g., the opportunity to help or congratulate a peer or the opportunity to avoid a confrontation with an aggressive student). Escape- or avoidance-motivated behaviors can be difficult to observe in natural settings because children rarely come into even remote contact with the feared stimuli. Sometimes the setting in which behaviors of interest occur make it difficult to collect direct observation data. For example, it may be difficult to observe child–parent interactions in home environments. In these instances, analogue assessment procedure may prove useful (Hintze, Stoner, & Bull, 2000).

Behavioral avoidance tests (BATs) have often been used to measure responses to feared or anxiety-producing stimuli (e.g., Van Hasselt, Hersen, Bellack, Rosenblum, & Lamparski, 1979). In some instances, children can be bought into proximity or direct contact with the specific stimuli and observers can record direct observation data (e.g., how close the child came to the dog and how long the child remained in the room with the dog). In addition, children can report their level of fear or anxiety while in these situations (Bellack, Kay, & Murrill, 1989). Sometimes, children are prompted to perform gradually more and more fear-provoking responses and observers can record how far the child progressed through the hierarchy (e.g., got within three feet of the dog, got within one foot of the dog, and petted the dog). Observers can also record the number or levels of prompts needed for the child to complete each step. For example, prompts could be less intrusive (e.g., a verbal prompt instructing the child to pet the dog), moderately intrusive (e.g., modeling petting the dog), or highly intrusive (e.g., use a hand-over-hand procedure to physically guide the child's hand over the dog's coat).

Because it is difficult to unobtrusively collect naturalistic observation data on parent–child interactions, analogue conditions are often useful for assessing these interactions. Forehand and McMahon (1981) developed an analogue procedure where children and parents took turns playing a game where each made up their own rules. During the parent game session, children received a series of commands and child (e.g., compliance) and parents responses were observed and recorded. Barkley (1997) devel-

oped similar direct observation procedures to collect data on child–parent interactions within an analogue setting (e.g., clinic). However, Barkley's procedure did not use a game format. Rather, parents were merely instructed to provide a series of scripted commands.

With some role-play measures children are provided descriptions of scenarios and are asked to imagine themselves in these situations and to respond as they would if they were actually in the described situation. Observers then use direct observation procedures to record responses of the children. One of the earlier role play measures was the Behavioral Assertiveness Test for Children (BAT-C) developed by Bornstein, Bellack, and Hersen (1997). With this measure, as children responded to described situations, observers recorded duration of eye contact, intensity of speech, requests for new behavior, and overall assertiveness. Elliott and Gresham (1991) developed social skills assessment scenarios for cooperation, assertion, responsibility, empathy, and self-control. After the situation is described, students provide information about the scenario verbally (e.g., define skill being prompted) and then are asked to demonstrate the skill (e.g., asking for help).

A variety of other analogue assessment procedures have been developed (Camp & Bash, 1981; Goldstein, 1999; Spivak, Platt, & Shure, 1976). With some procedures children make overt responses to contrived stimuli. In other situations, students make overt responses to described situations or scenarios. With other analogue assessment procedures contrived situations may be presented and students provide verbal or written descriptions on how they may or should have responded (e.g., Goldstein, 1999). Finally, in some instances both the situation and the response are described (e.g., Spivak et al., 1976).

Experimental Functional Analysis Procedures

Experimental functional analysis procedures also employ analogue assessment conditions as well as direct observation and empirical recording of target behaviors. These procedures differ from previously described analogue assessment procedures in that they are not designed to identify or confirm problems. Rather, single-subject experimental design methodology is used to attempt to determine the function of problem behaviors after they have been identified.

Experimental functional analysis procedures are based on operant behavioral psychology. Under this model, behaviors are maintained through positive or negative reinforcement. Furthermore, these positive and negative reinforcers can be idiosyncratic. Thus, two children may present similar behavior problems. These two children may engage in topographically similar behavior (e.g., both child repeat what others say), at similar rates, during similar general environmental conditions (e.g., during school only). However, one child may engage in the behavior in order to receive attention (positive reinforcement) and the other may engage in the behavior to avoid working on a specific task (negative reinforcement or escape/avoidance behavior). By determining the function of the behavior, experimental functional analysis procedures not only suggest classes of interventions that may be effective but may also prevent one from developing and implementing interventions that may exacerbate behavior problems. For example, ignoring a child who is tantrumming is likely to strengthen the behavior if its function is to escape or avoid attention.

With experimental functional analysis procedures, students are placed in analogue environments or conditions for brief periods. These conditions are designed to simulate a child's natural environment. However, these conditions are much more tightly controlled and only one variable at a time is intentionally altered. These experimental procedures allow for comparisons of target behavior levels across conditions in order to determine the variables that may be maintaining target behaviors in the child's natural environment.

Typically, children are exposed to conditions that test positive and negative reinforcement hypotheses. In addition, children may be exposed to a control condition. Under the positive reinforcement conditions, the child is placed in an environment and allowed to engage in preferred behavior. Reinforcement is delivered following the occurrence of the target behavior. Typically reinforcers tested include attention (e.g.,

parent or teacher approach and interact with the child only after the child engages in target behaviors) and access to tangibles (e.g., students are given preferred toys for a brief period only after engaging in target behaviors). After a brief period, reinforcers are removed (e.g., parent takes preferred item or teacher turns away from child) and the condition continues.

Under the negative reinforcement condition, the child is given a task or demand and that task or demand is removed only when the child engages in the target behavior. After a specified amount of time (usually a brief period, e.g., 15–30 seconds) the child is once again presented with the task or demand. Often a control condition is implemented during which students are given access to preferred items or activities and attention, but delivery or removal of these items, activities, and attention is never delivered contingent upon the child's target behavior. This condition is designed to determine if performing the target behavior is reinforcing in and of itself (e.g., self-stimulation or self-reinforcement).

Typically, children are exposed to the various conditions during brief sessions (e.g., 10–20 minutes) where only one condition is tested. After a brief break, other conditions follow. Session times are equivalent and conditions are presented in random or counterbalanced order. Although frequency counts are often used, observers could use any of the empirical data collection procedures to record levels of target behaviors for each session.

The number of exposures to conditions required may vary (see Cooper, Wacker, Sasso, Reimers, & Donn, 1990, for examples of brief experimental functional analysis procedures and Iwata, Vollmer, & Zarcone, 1990, for examples of longer procedures). However, the goal is to collect data until clear levels and trends occur within and across conditions. When this occurs, comparing target behavior levels across conditions can indicate the function or functions of the target behavior. Once the function of a behavior is known, treatments can be designed and implemented within the child's natural environment that are based on assessed function(s).

Experimental functional analysis procedures have been used to identify the function(s) of a variety of behaviors (e.g., aggression, loud vocalizations, noncompliance, self-injurious behavior, wandering, and echolalia) across settings and children (Crawford, Brockel, Schauss, & Miltenberger, 1992; Derby et al., 1992; Iwata et al., 1982; Mace, Webb, Sharkey, Mattson, & Rosen, 1988; Sasso et al., 1992; Townsend, 2000). It is beyond the scope of this chapter to review all the possible variations across experimental conditions, observations and recording procedures, and experimental design elements (see McComas et al., 2000, for a review of these procedures). Furthermore, although functional analysis procedures may indicate some general treatment strategies (e.g., extinction and differential reinforcement of alternative behaviors) because treatment selection, development, and implementation are likely to be influenced by other variables, including available resources, expertise of teacher or parent, and time constraints, it is not feasible to review all possible treatment alternatives suggested by different functional analysis outcomes.

Consideration When Using Analogue Assessment Procedures

Analogue assessment procedures allow one to collect direct observation data that may be difficult to collect in natural environments. However, there are several limitations associated with experimental functional analysis and all other analogue assessment procedures. First, exposing children to analogue conditions can cause distress in children (Townsend, 2000). During experimental functional analysis procedures, conditions are established that result in inappropriate behaviors being reinforced. In some situations it may not be appropriate to occasion or reinforce dangerous or disruptive behaviors (Iwata et al., 1990; McComas et al., 2000). In addition, experimental functional analysis procedures may not be useful when problem behaviors occur at low rates (e.g., temper tantrums).

Perhaps the most serious concern with analogue assessment procedures is that the artificial conditions may not be sufficiently representative of the child's natural environment. With many analogue assessment procedures, children are placed in novel, unfa-

miliar situations and often assessment procedures are conducted by someone unfamiliar with the child (e.g., a behavior analyst or school psychologist as opposed to the child's teacher or parent). Thus, children's behavior under these conditions may not represent their behavior in their natural environment (Lennox & Miltenberger, 1989; Sasso et al., 1992). Running experimental functional analysis procedures in children's natural environments may enhance the continuity between analogue and natural conditions (Sasso & Reimers, 1988; Watson, Ray, Sterling, & Logan, 1999).

With experimental functional analysis procedures, children are exposed to multiple analogue conditions. It can be difficult to match multiple conditions with natural environments. For example, one could fail to identify preferred stimuli to be used as reinforcers or demand conditions that are appropriate for testing the escape/avoidance function. Even when stimuli from the natural environment are accurately identified, contingent delivery or removal of these stimuli may operate differently on target behaviors across natural and analogue conditions (Iwata et al., 1990). Reinforcement rates and immediacy of reinforcement during experimental functional analysis procedures may not represent those in natural environments. Both these variables can have an impact on behavior. Finally, experimental functional analysis procedures employ highly sensitive measures that may be more susceptible to confounding variables (e.g., hunger, nervousness, habituation, and fatigue).

When analogue conditions ask children to imagine themselves in a specific situation, the ability of children to respond to described scenarios is likely to be highly variable both (1) within conditions, across children, and (2) within children, across conditions. Regardless, the quality of the child's internal behavior (imagining) cannot be measured and is likely to affect the child's response. Similarly, when children are asked to describe their responses, they may not provide responses that reflect how they would actually respond. Instead, demand characteristics associated with assessment conditions may cause children to respond how they think they should respond or how they think the person assessing them would want them to respond. These limitations suggest that any hypothesis formed via analogue assessment procedures may need to be confirmed by collecting data on the children's behavior within their natural environments.

SUMMARY AND CONCLUSIONS

The direct nature of the behavioral assessment procedures reviewed is often appealing to those who are uncomfortable with inferring within-subject problems based on overt behaviors and making inferences regarding the cause of those problems. Even though the assessment procedures described and analyzed here are more direct than traditional procedures, these procedures can still yield data that do not accurately reflect a child's behavior in his or her natural environment. Therefore, in most instances it is useful to confirm data collected via all the procedures described previously. This confirmatory data can be collected using other direct assessment procedures. For example, analogue assessment data can be confirmed by collecting direct observation data in a child's natural environment, interviewing a teacher, parent, or peer (e.g., Bergan & Kratochwill, 1990) or with checklist and rating scales (e.g., Achenbach, 1991). Data may also be confirmed by using multiple observers (see House, House, & Campbell, 1981, for procedures and formula for confirmed direct observation data).

Confirming data before interventions are implemented may enhance the confidence one has in conclusions that have been based on those data. However, the ultimate purpose of collecting data is to provide information that will lead to more effective interventions or education programs. By indicating the strengths and limitations associated with direct behavioral assessment procedures, it is hoped that readers will be more likely to develop or choose assessment procedures that will yield data that accurately represent problem behaviors and also allow for the identification of variables that may be maintaining those target behaviors. However, because there are limitations associated with all direct assessment procedures, it is essential that assessment processes do not halt after problems have been identified

or interventions have been developed based on hypothesized functions of behavior. Rather, one must continue to assess to determine if the interventions that were developed based on these assessment data were effective in bringing about socially significant changes in the target behavior(s) within the child's natural environment (Barnett et al., 2000).

REFERENCES

Achenbach, T. M. (1991). *Manual for the Child Behavior Checklist/4-18 and 1991 profile*. Burlington: University of Vermont, Department of Psychology.

Alessi, G., & Kaye, J. H. (1983). *Behavior assessment for school psychologists*. Kent, OH: National Association of School Psychologists.

Barkley, R. A. (1997). *Defiant children: A clinician's manual for assessment and parent training* (2nd ed.). New York: Guilford Press.

Barnett, D. W., Daley, E. J., Hampshire, E. M., Hines, N. R., Maples, K. A., Ostron, J. K., & Van Buren, A. E. (1999). Meeting performance-based training demands: Accountability in an intervention based practicum. *School Psychology Quarterly, 14,* 357–379.

Bellack, L., Kay, S. R., & Murrill, L. M. (1989). The Dysphorimeter: An objective analogue for the assessment of depression, anxiety, pair and other dysphoric states. *American Journal of Psychotherapy, 43,* 260–268.

Bergan, J. R., & Kratochwill, T. R. (1990). *Behavioral consultation in applied settings*. New York: Plenum Press.

Bornstein, M. R., Bellack A. S., & Hersen, M. (1977). Social skills training for unassertive children: A multiple-baseline analysis. *Journal of Applied Behavior Analysis, 10,* 183–195.

Bowers, F. E., McGinnis, C., Ervin, R. A., & Friman, P. C. (1999). Merging research and practice: The example of positive peer reporting applied to social rejection. *Education and Treatment of Children, 22,* 218–226.

Camp, B. W., & Bash, M. A. (1981). Think aloud. Champaign, IL: Research Press.

Carded-Smith, L. K., & Fowler, S. A. (1984). Positive peer pressure: The effects of peer monitoring on children's disruptive behavior. *Journal of Applied Behavior Analysis, 17,* 213–227.

Carr, E. G. (1993). Behavior analysis is not ultimately about behavior. *The Behavior Analyst, 16,* 47–49.

Cashwell, C. S, Skinner, C. H., Dunn, M., & Lewis, J. (1998). Group reward programs: A humanistic approach. *Humanistic Education and Development, 37,* 47–53.

Cole, C. L, Marder, T., & McCann, L. (2000). Self-monitoring. In E. S. Shapiro & T. R. Kratochwill (Eds.), *Conducting school-based assessment of child and adolescent behavior* (pp. 121–149). New York: Guilford Press.

Cone, J. D. (1978). The behavioral assessment grid (BAG): A conceptual framework and a taxonomy. *Behavior Therapy, 9,* 882–888.

Cooper, L., Wacker, D., Sasso, G., Reimers, T., & Donn, L. (1990). Using parents as therapists to evaluate appropriate behavior of their children: Application to a tertiary diagnostic clinic. *Journal of Applied Behavior Analysis, 23,* 285–296.

Crawford, J., Brockel, B., Schauss, S., & Miltenberger, R. G. (1992). A comparison of methods for the functional assessment of stereotypic behavior. *Journal of the Association for Persons with Severe Handicaps, 17,* 77–86.

Derby, M., Wacker, D., Sasso, G., Steege, M., Northup, J., Cigrand, K., & Asmus, J. (1992). Brief functional assessment techniques to evaluate aberrant behavior in an outpatient setting: A summary of 79 cases. *Journal of Applied Behavior Analysis, 25,* 713–721.

Elliott, S.N., & Gresham, F. M. (1991). *Social skills intervention guide: Practical strategies for social skills training*. Circle Pines, MN: American Guidance Services.

Ervin, R. A., Johnston, E. S., & Friman, P. C. (1998). Positive peer reporting to improve the social interactions of a socially rejected girl. *Proven practice: Prevention and remediation solutions for school problems, 1,* 17–21.

Ervin, R. A., Miller, P. M., & Friman, P. C. (1996) Feed the hungry bee: Using positive peer reports to improve the social interactions and acceptance of a socially rejected girl in a residential placement. *Journal of Applied Behavior Analysis, 29,* 251–253.

Fixsen, D. L., Phillips, E. L., & Wolf, M. M. (1972). Achievement Place: The reliability of self-reporting and peer-tutoring and their effects on behavior. *Journal of Applied Behavior Analysis, 5,* 19–30.

Forehand, R. L., & McMahon, R. J. (1981). *Helping the noncompliant child: A clinician's guide to parent training*. New York: Guilford Press.

Glynn, E. L., & Thomas, J. D. (1974). Effects of cueing on self-control of classroom behavior. *Journal of Applied Behavior Analysis, 7,* 299–306.

Goldstein, A. P. (1999). *The prepare curriculum* (rev ed.). Champaign, IL: Research Press.

Greenwood, C. R., Dinwiddie, G., Terry, B., Wade, L., Stanely, S. O., Thibadeau, S., & Delquadri, J. C. (1987). Teacher- versus peer-mediated instruction: An eco-behavioral analysis of achievement outcomes. *Journal of Applied Behavior Analysis, 17,* 521–538.

Gresham, F. M. (1998). Designs for evaluating behavior change: Conceptual principles of single-case methodology. In T. S. Watson & F. M. Gresham (Eds.), *Handbook of child behavior therapy* (pp. 23–40). New York: Plenum Press.

Hartmann, D. P., Roper, B. L., & Bradford, C. C. (1979). Some relationships between behavioral and traditional assessment. *Journal of Behavioral Assessment, 1,* 3–21.

Hayes, S. C., & Cavoir, N. (1977). Multiple tracking and the reactivity of self-monitoring: I. Negative behaviors. *Behavior Therapy, 8,* 819–831.

Hendrickson, J. M., Gable, R. A., Conroy, M. A., Fox,

J., & Smith, C. (1999). Behavior problems in schools: Ways to encourage functional behavior assessment (FBA) of discipline-evoking behavior of students with emotional and/or behavioral disorders (EBD). *Education and Treatment of Children, 22,* 280–290.

Henington, C., & Skinner, C. H. (1998). Peer monitoring. In K. Toppings & S. Ely (Eds.), *Peer assisted learning* (pp. 237–253). Hillsdale, NJ: Erlbaum.

Hintze, J. M., & Shapiro, E. S. (1995). Best practices in systematic observation of classroom behavior. In A. Thomas & J. Grimes (Eds.), *Best practices in school psychology-III* (pp. 651–660). Washington, DC: National Association of School Psychologists.

Hintze, J. M., Stoner, G., & Bull, M. H. (2000). Analogue assessment: Emotional/behavior problems. In E. S. Shapiro & T. R. Kratochwill (Eds.) *Conducting school-based assessment of child and adolescent behavior* (pp. 55–77). New York: Guilford Press.

House, A. E., House, B. J., & Campbell, M. B. (1981). Measures of interobserver agreement: Calculation formulas and distribution effects. *Journal of Behavioral Assessment, 3,* 37–58.

Iwata, B. A., Dorsey, M. F., Sifler, K. J., & Richman, G. S. (1982). toward a functional analysis of self-injury. *Analysis and Intervention in Developmental Disabilities, 3,* 1–20.

Iwata, B. A., Vollmer, T. R., & Zarcone, J. R. (1990). The experimental (functional) analysis of behavior disorders: Methodology, applications, and limitations., In A. C. Repp & N. N. Sigh (Eds.), *Perspective on the use of nonaversive and aversive interventions for persons with developmental disabilities* (pp. 301–330). Sycamore, IL: Sycamore.

Johnson, S. M., & Bolstad, O. D. (1973). Methodological issues in naturalistic observation: Some problems and solutions. In L. A. Hamerlynch, L. E. Handy, & E. J. Marsh (Eds.), *Behavior change: Methodology, concepts, and practices* (pp. 7–68). Champaign, IL: Research Press.

Jones, K. M., Young, M. M., & Friman, P. C. (2000). Increasing peer praise of socially rejected delinquent youth: Effects on cooperation and acceptance. *School Psychology Quarterly, 15,* 30–39.

Kazdin, A. E. (1977). Artifact, bias, and complexity of assessment: The ABC's of reliability. *Journal of Applied Behavior Analysis, 10,* 141–150.

Kern, L., Marder, T. J., Boyajian, A. E., Elliot, C. M., & McElhattan, D. (1997). Augmenting the independence of self-management procedures by teaching self-initiation across settings and activities. *School Psychology Quarterly, 12,* 23–32.

Lalli, J. S., Browder, D. M., Mace, F. C., & Brown, K. (1993). Teacher use of descriptive analysis data to implement interventions to decrease students' maladaptive behavior. *Journal of Applied Behavior Analysis, 10,* 141–150.

Lennox, O. I., & Miltenberger, R. G. (1989). Conducting a function assessment of problem behavior in applied settings. *Journal of the Association for Persons with Severe Handicaps, 14,* 304–311.

Lentz, F. E. (1982). *An empirical examination of the utility of partial interval and momentary time sampling as a measure of behavior*. Unpublished doctoral dissertation, The University of Tennessee, Knoxville.

Lloyd, M. E., & Hilliard, A. M. (1989). Accuracy of self-recording as a function of repeated experience with different self-control contingencies. *Child and Family Behavior Therapy, 11,* 1–14.

Mace, F. C., Webb, M. E., Sharkey, R. W., Mattson, D. M., & Rosen, H. S. (1988). Functional analysis and treatment of bizarre speech. *Journal of Behavior Therapy and Experimental Psychology, 19,* 714–721.

McComas, J. J., Hoch, H., & Mace, F. C. (2000). Functional analysis. In E. S. Shapiro & T. R. Kratochwill (Eds.), *Conducting school-based assessment of child and adolescent behavior* (pp. 78–101). New York: Guilford Press.

McCurdy, B. L., & Shapiro, E. S. (1992). A comparison of teacher-, peer-, and self-monitoring with curriculum-based measurement in reading among students with learning disabilities. *Journal of Special Education, 26,* 162–180.

McLaughlin, T. F., & Skinner, C. H. (1996). Improving academic performance through self-management: cover, copy, and compare. *Intervention in School and Clinic, 32,* 113–118.

Myerson, J., & Hale, S. (1984). Practical implications of the matching law. *Journal of Applied Behavior Analysis, 17,* 367–380.

Nelson, R. O. (1977). Methodological issues in assessment via self-monitoring. In J. D. Cone & R. P. Hawkins (Eds.), *Behavioral assessment: New directions in clinical psychology* (pp. 217–240). New York: Brunner/Mazel.

Nelson, R. O., Hay, L. R., Devany, J., & Koslow-Green, L. (1980). The reactivity of children's self-monitoring: Three experiments. *Child Behavior Therapy, 2,* 1–24.

Nelson, R. O., Hay, L. R., Hay, W. M., & Carsten, C. B. (1977). The reactivity and accuracy of teachers' self-monitoring of positive and negative classroom verbalizations. *Behavior Therapy, 8,* 972–985.

Nelson, R. O., & Hayes, S. C. (1979). Some current dimensions of behavioral assessment. *Behavioral Assessment, 1,* 1–16.

Nelson, R. O., & Hayes, S. C. (Eds.). (1986). *Conceptual foundations of behavioral assessment.* New York: Guilford Press.

Nelson, J. R., Roberts, M. L., & Rutherford, R. B., Mathur, S. R., & Aaroe, L. A. (1999). A statewide survey of special education administrators and school psychologists regarding functional behavioral assessment. *Education and Treatment of Children, 22,* 267–279.

O'Neill, R. E., Horner, R. H., Albin, R. W., Sprague, J. R., Storey, K., & Newton, J. S. (1997). *Functional assessment of problem behavior: A practical assessment guide* (2nd ed.). Pacific Grove, CA: Brooks/Cole.

Powell, J., Martindale, B., Kulp, S., Martindale, A., & Bauman, R. (1977). Taking a closer look: time sampling and measurement error. *Journal of Applied Behavior Analysis, 10,* 325–332.

Reschly, D. J., & Ysseldyke, J. E. (1995). School psychology paradigm shifts. In A. Thomas & J. Grimes (Eds.), *Best practices in school psychology-III* (pp.

17–31). Washington, DC: National Association of School Psychologists.

Robinson, S. L. (1998). *Effects of positive statements made by peers on peer interactions and social status of children in a residential treatment setting*. Unpublished doctoral dissertation, Mississippi State University, Starkville.

Santogrossi, D. A. (1974, October). *Self-reinforcement and external monitoring of performance on an academic task*. Paper presented at the fifth annual conference on Applied Behavior Analysis in Education, Kansas City, KS.

Sasso, G. M., & Reimers, T. M. (1988). Assessing the functional properties of behavior: Implications and applications for the classroom. *Focus on Autistic Behavior, 3,* 1–15.

Sasso, G. M., Reimers, T. M., Cooper, L. J., Wacker, D., Berg, W., Steege, M., Kelly, L., & Allaire, A. (1992). Use of descriptive and experimental analyses to identify the functional properties of aberrant behavior in school settings. *Journal of Applied Behavior Analysis, 25,* 809–821.

Saudargus, R. A. (1992). *State-Event Classroom Observation system (SECOS)*. Knoxville: University of Tennessee, Department of Psychology.

Shapiro, E. S. (1984). Self-monitoring. In T. H. Ollendick & M. Hersen (Eds.), *Child behavior assessment: Principals and procedures* (pp. 148–165). New York: Pergamon Press.

Shapiro, E. S. (1987). *Behavioral assessment in school psychology*. Hillsdale, NJ: Erlbaum.

Shapiro, E. S. (1996). *Academic skills problems: Direct assessment and intervention* (2nd ed). New York: Guilford Press.

Shapiro, E. S., & Browder, D. M. (1990). Behavioral assessment: Applications for persons with mental retardation. In J. Matson (Ed.), *Handbook of behavior modification with persons with mental retardation* (2nd. ed., pp. 93–122). New York: Plenum Press.

Shapiro, E. S., & Kratochwill, T. R. (2000). Introduction: Conducting a multidimensional behavioral assessment. In E. S. Shapiro & T. R. Kratochwill (Eds.), *Conducting school-based assessments of child and adolescent behavior* (pp.1–20). New York: Guilford Press.

Shapiro, E. S., McGonigle, J. J., & Ollendick, T. H. (1980). An analysis of self-assessment and self-reinforcement in a self-managed token economy with mentally retarded children. *Applied Research in Mental Retardation, 1,* 227–240.

Shapiro, E. S., & Skinner, C. H. (1990). Best practices in observational/ecological assessment. In A. Thomas & J. Grimes (Eds.), *Best practices in school psychology-II* (pp. 17–31). Washington, DC: National Association of School Psychologists.

Skinner, C. H., Cashwell, T. H., & Skinner, A. L. (2000). Increasing tootling: The effects of a peer monitored group contingency program on students' reports of peers' prosocial behaviors. *Psychology in the Schools, 37,* 263–270.

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 practice* (2nd ed., pp. 19–45). New York: Guilford Press.

Skinner, C. H., Rhymer, K. N., & McDaniel, E. C. (2000). Naturalistic direct observation in educational settings. In E. S. Shapiro & T. R. Kratochwill (Eds.), *Conducting school-based assessments of child and adolescent behavior* (pp. 21–54). New York: Guilford Press.

Skinner, C. H., Shapiro, E. S., Turco, T. L., Cole, C. L., & Brown, D. K. (1992). A comparison of self- and peer delivered immediate corrective feedback on multiplication performance. *Journal of School Psychology, 30,* 101–116.

Skinner, C. H., & Smith, E. S. (1992). Issues surrounding the use of self-management interventions for increasing academic performance. *School Psychology Review, 21,* 202–210.

Spivak, G., Platt, J., & Shure, M. B. (1976). *The problem-solving approach to adjustment*. San Francisco: Jossey-Bass.

Stern, G. W., Fowler, S. A., & Kohler, F. W. (1988). A comparison of two intervention roles: Peer monitor and point earner. *Journal of Applied Behavior Analysis, 21,* 103–109.

Telzrow, C. F. (1999). IDEA amendments of 1997: Promise or pitfall for special education reform? *Journal of School Psychology, 37,* 7–28.

Townsend, B. J. K. (2000). *A functional analysis of the echolalic behavior of three children with autism in a residential school setting*. Unpublished doctoral dissertation, Mississippi State University, Starkville.

Van Hasselt, V. B., Hersen, M., Bellack, A. S., Rosenblum, N. D., & Lamparski, D. (1979). Tripartite assessment of the effects of systematic desensitization in a multi-phobic child: An experimental analysis. *Journal of Behavior Therapy and Experimental Psychiatry, 10,* 51–55.

Watson, T. S., Ray, K. R., Sterling, H. E., & Logan, P. (1999). Teacher-implemented functional analysis and treatment: A method of linking assessment to intervention. *School Psychology Review, 28,* 292–302.

Yell, M. L., & Shriner, J. G. (1997). The IDEA amendments of 1997: Implications for special and general education teachers, administrators, and teacher trainers. *Focus on Exceptional Children, 30,* 1–19.