



Behavioral Assessment In Schools

Nasir Mohammad Bhat

Asst. Professor , Dept. of Education,

Central University of Kashmir, 190004 - Srinagar, J&K, India

E-mail:nasirbhat111@gmail.com

Abstract: *The assessment of social and emotional functioning in the schools is needed for purposes of classification, intervention development, and measurement of outcomes. Procedures for completing these evaluations, as well as the relative strengths and weaknesses of the techniques, are described in this article. The descriptions are organized by purpose, beginning with techniques for problem identification and followed by the application of these techniques to measuring outcomes. Legal, ethical, and cultural issues as they relate to these procedures are described at the end.*

Key words: *Traditional Assessment, Analogue Assessment, Functional Behavioral Assessment*

INTRODUCTION:

The evaluation of behavioral and emotional problems in the classroom is warranted by the difficulties experienced by children who contend with these problems on a day-to-day basis. A number of behavioral and emotional problems that can prevent a child from functioning at a level commensurate with his or her same age or same-grade peers are commonly evidenced in the school setting. Some of the more common behavioral and emotional problems that cause such impairments include anxiety, attention deficit/hyperactivity disorder (ADHD), anger/aggression, attachment disorders, conduct problems, depression, delinquency, posttraumatic stress disorder (PTSD), and bullying. The school environment is unique in that it places demands on children that are not otherwise required of them. For example, in the school setting, children are expected to sit quietly and attend to classroom activities. Furthermore, schools require students to function within demanding social contexts that youth may avoid or not be exposed to outside of school. The difficulties these expectations pose for children with behavioral and emotional problems are substantial, and assessment of these problems in

the setting where they often occur has important implications for the development of effective

interventions. It logically follows that assessment should occur, at least in part, in the school setting.

Evaluation of behavioral and emotional problems in schools is frequently done to determine whether a child or an adolescent is eligible for, or would benefit from, special services. Under these circumstances, evaluation is done to answer questions pertaining to classification and to guide the development of effective interventions. Although determining eligibility for special services is important, it is the development of interventions that is vital for changing problem behaviors. Effective interventions are those that target the presenting behaviors and not the label. Therefore, school professionals should place an emphasis on the development and implementation of interventions that target the specific needs of the students who are referred for evaluation rather than their diagnoses or labels.

Traditional Assessment

There are a number of assessment methods used to evaluate behavioral and emotional problems in the schools. Traditional assessment involves the examination of scores from norm-referenced

measures to determine a particular individual's difference from or similarity to a normative sample. In other words, traditional assessment asks the question, "How does this student's behavior compare with that of similar children?" Standard tests of cognitive abilities and academic achievement, as well as many rating scales, are structured in this way. Using traditional assessment techniques, the inferred construct being measured (e.g., attention, depression) is thought of as a cause of a particular behavior in a specific setting. Although identifying the underlying syndrome or "label" related to a child's behavior is certainly important for diagnosis and classification, this information is not always the most useful for designing interventions. In contrast, measurement approaches such as behavioral observation, analogue assessment, interviews, and rating scales provide highly useful information about the functional relationship between the environment and particular behaviors. These techniques attempt to uncover the causes and consequences of behavior that often lead directly to effective intervention strategies.

Direct Observation

Direct observation is likely the most straightforward of all assessment techniques because it involves just what it implies: the direct observation and recording of problem behaviors in the context in which they occur. Three methods of direct observation are narrative recording (i.e., writing down what is happening in a narrative fashion, often following the antecedent-behavior-consequence [ABC] pattern or what occurs before, during, and after specific behaviors), event recording (i.e., tallying the instance of a target behavior), and time sampling (i.e., breaking down observations into intervals and recording the number of times a behavior occurs relative to the number of intervals observed to form a ratio). Direct observation differs from other methods of assessment in that little inference is required, making it one of the less biased methods for evaluating potential target behaviors.

There are some measurement issues related to observational methods that can negatively influence the data that are collected, including poorly trained observers, observer drift, poor operational

definitions of target behaviors, poorly constructed data-recording systems, and reactivity (i.e., when the act of observation affects the student's behavior). Establishing inter observer agreement is one way in which to ensure that the data collected from direct observation are reliable and useful. A common procedure used to measure and enhance the quality of direct observation data is having another trained observer collect data simultaneously with the primary observer and then comparing data across observers to yield a measure of inter observer agreement.

Direct observation is a labor intensive procedure that requires significant training time and effort to maintain reliability. An additional problem with direct observation is the limited availability of the technique to assess infrequently occurring behavior. For example, children who are perceived as aggressive by teachers might exhibit only two or three episodes of these behaviors per week (or even fewer). The likelihood of observing these defiant or aggressive behaviors during discrete classroom observations is small. The technique is probably most practical for high-frequency behaviors such as on-task behavior and "talking out." Finally, most observational data are difficult to put into context because there are no norms. In the school environment there is considerable variability in behavioral expectations across classroom activities and between teachers, resulting in various and shifting definitions of what is "normal." There are observational techniques that allow one to begin to account for this shifting context. For example, some time-sampling procedures require the observer to shift back and forth between observing the target child and same-sex classmates or a randomly selected peer. Pooling the data from the target and peer intervals provides the evaluator with a definition of what was normal during the context observed. However, no strategy will always work to overcome the problem of shifting contexts, and school professionals must interpret the results of their observations carefully.

Analogue Assessment

Analogue assessment provides an alternative to direct observation and is useful in situations where direct observation is not feasible. Analogue assessment is an observational measure of the

child's behavior of interest elicited through the use of simulated conditions; observations are made in the simulated setting to predict behavior in real-life settings. Assessments of this type can be administered by paper and pencil, by audiotape or videotape, or through the use of role-plays or enactments. In a controlled environment, the evaluator creates a situation parallel to one that the child experiences in school and observes and records the behavior of interest. As with direct observation methods, this technique can be labor intensive; however, it does allow the evaluator to assess specific conditions that may occur only rarely. Some of these situations may be easy to recreate, for example, asking a child who is poor in math to complete some math and reading work. In this example, differences in the child's reactions to the various academic tasks may be observed, and these data may contribute to an understanding of behavior problems experienced during math class. Among the problems associated with this method of assessment is the lack of standardization, issues related to reliability and validity, and problems with generalization. The most basic of these is the fact that analogue assessments are contrived situations, and regardless of how similar the situation is to the targeted environment, the child usually knows that it is contrived, and this knowledge may influence his or her behavior. Whenever analogue assessment techniques are used, all of these potential issues should be considered when interpreting the results.

Interviews

Interviews to assess behavioral and emotional problems in children can be conducted with the child himself or herself or with the child's parents, teachers, and/or other caregivers. A variety of different interview formats exist. An omnibus interview is designed for gathering a wide range of information, whereas a behavior-specific interview is narrower in scope and focuses on the assessment of specific problem areas. An interview that is problem solving in nature focuses on presenting concerns with the goal of developing an intervention plan. Interviews can be structured, unstructured, or semi structured. Although it is often the case that interviews focus on reaching a diagnostic conclusion, interviews

that focus on both problem assessment and the development of intervention plans are useful in school settings. Behavioral interviewing addresses the assessment of the current behaviors and environmental factors that contribute to the initiation and maintenance of problem behaviors. These interviews are designed to detect specific problems that can be targeted for intervention and include phases of problem identification, problem analysis, treatment implementation, and treatment evaluation. Interviews may be a valuable and efficient method for collecting information about a wide range of behavioral and emotional problems. The evaluator can efficiently collect data about behavior over a long period of time and in a variety of situations. In addition, the adults being interviewed may provide information about their reactions to this behavior or interventions related to the targeted behavior. Ideally, interviews are conducted by staff members who will continue to be involved in developing and implementing interventions with students in need. In these cases, the interview can be used to establish the initial relationship and communicate empathy and support to the child while collecting valuable information.

Rating Scales

Rating scale measures provide a standardized format to gather information about an individual's behavioral characteristics and can be administered to children to obtain self-report data or to parents, teachers, and/or other caregivers to obtain informant-report data. A few of the advantages of rating scales are their ability to provide behavioral information in a short amount of time, their moderate cost, and their technical precision and practical utility. Rating scales can be used in the comprehensive assessment of potential behavior problems and deficits in functioning for screening, diagnosis/classification, treatment selection/intervention planning, progress monitoring, and outcome evaluation.

As with most other forms of assessment, there are some problems associated with the use of rating scales. These problems typically fall under one of two categories: response bias and error variance. Response bias refers to trends in the way in which informants respond to rating scales. Some examples of response bias are halo effects (i.e.,

rating the student in a positive or negative manner based on characteristics unrelated to what is being measured), leniency or severity effects (i.e., the tendency to have an overly generous or overly critical response set), and central tendency effects (i.e., the tendency to rate down the middle and avoid end point ratings). Error variance refers to the differences among several ratings of the same individual. The different types of error variance include source variance (i.e., differences in ratings provided by various persons), setting variance (i.e., differences in ratings across situations), temporal variance (i.e., inconsistency of behavior ratings over time), and instrument variance (i.e., differences in ratings using different instruments).

Although rating scales are valuable tools, it is frequently tempting to over rely on them due to their minimal costs. Although they are efficient, there are many questions they do not answer (e.g., the role of environmental contingencies). Although environmental concerns are critical to the development of interventions, this information is generally not provided by rating scales. Furthermore, rating scales are only as useful as the sources. For example, children are fairly poor raters of their own disruptive, off-task, and inappropriate behaviors; consequently, adults have been identified as a better source for this information. On the other hand, children appear to be better at reporting their internal affective states and, as a result, are generally better raters of their own depressive and anxious characteristics. However, although these are general trends that should be considered when interpreting rating scales, there is considerable individual variability.

Functional Behavioral Assessment

A primary goal of FBA is to identify environmental conditions that are correlated with the occurrence and nonoccurrence of behaviors. This type of assessment provides a systematic analytic method for identifying those things that serve to initiate and maintain behavior. FBA incorporates many methods of assessment to gain a complete understanding of an individual's behavior in a particular environmental context. Approaches to FBA can be indirect (i.e., gathering behavioral information from rating scales, interviews, etc.), direct (i.e., directly observing the behavior in the context of the natural environment

in which it occurs), or experimental (i.e., generating hypotheses about the maintaining contingencies for the problem behavior and testing them with the individual serving as his or her own control).

Completing these assessments is an iterative process of data collection, hypotheses generation, and collection of additional data. The hypotheses provide the specific questions that guide the subsequent data collection that either confirms or redirects the hypotheses. The identification of the cause-effect relations in the interaction of an individual and his or her environment is integral to the development of interventions aimed at changing behavior. Because FBA focuses on the relationship between particular environmental circumstances the development of effective interventions for problem behaviors. A successful FBA provides information about these environmental circumstances affecting the problem behavior and also provides a clear direction for treatment.

CURRENT ISSUES

Because all of the assessment techniques discussed present their own unique advantages and limitations, the use of multiple assessment techniques with multiple informants is often considered the best method for evaluating behavioral and emotional problems. Although the use of multiple informants (e.g., parents, teachers, students) has the potential to provide a wealth of information, inconsistencies in data from one source and method to another is an issue that is often encountered in behavioral and emotional assessment. Such inconsistencies must be taken into consideration when assessing behavioral and emotional problems and especially when identifying behaviors to target for intervention. Frequently, careful evaluation relying on multiple informants using multiple methods will yield data that are contradictory and inconsistent, making interpretation difficult. Nevertheless, this comprehensive strategy is recommended to reduce the influence of specific sources of error and to produce the opportunity to recognize and interpret convergent themes used to generate explanations and interventions.

EVALUATING OUTCOMES

As noted in the previous section, children and adolescents manifest many of their problems related to behavioral and emotional problems at school. This not only warrants school-based evaluation of these problems, it frequently results in school-based provision of services. These services take many forms, ranging from universal interventions, to disorder-specific treatment, to schoolbased day treatment programs. Professionals providing these services are expected to assess the progress of children and adolescents participating in them so as to modify interventions and know when interventions are no longer needed. As a result, the questions that guide outcome evaluations pertain to understanding an individual's progress in treatment with implications for modifying or discontinuing services. These assessment procedures frequently include observations, rating scales, and performance measures.

Observations

Observations have inherent advantages and disadvantages that are described in the previous section. The use of observation measures to measure outcomes is warranted with high-frequency behavior or with behavior that has a high degree of situational specificity. Observed rates of on-task behavior is a common outcome measure of a high-frequency behavior. Defiant and disruptive behavior is usually difficult to evaluate with observational techniques because its frequency is low even for students who teachers perceive as defiant or disruptive. This is especially true in secondary schools, where episodes of defiant behavior can be quite intimidating and need not occur very often for a teacher to report a problem. Low-frequency behavior with a high degree of situational specificity is conducive to observational measures because the observer can efficiently target the specific situation that elicits the problematic behavior. For example, a child referred for problematic behavior in the cafeteria may warrant observational assessment because the observation period is limited and the likelihood of observing the problematic behavior within this situation may be quite high. Observational measures have been found to be sensitive to intervention effects when administered frequently and reliably.

Rating Scales

The use of rating scales can be a very efficient method for collecting outcome data. Compromises to validity can occur when raters are invested in the intervention process, resulting in a bias to perceive improvement. Collecting data from some raters who are blind to treatment or not invested in its success can be a valuable assessment strategy; however, sometimes these people are not available or are not able to provide useful assessment data. Collecting rating scale data from multiple sources and then interpreting convergent findings has been recommended so as to minimize individual biases. This technique can be helpful, but it also presents evaluators with unique challenges. It is quite common to receive rating scale data with tremendous variability and little convergence around any conclusion. Although this may indicate poor reliability, it may also indicate that the child's behavior changes as a function of classroom, activity, or teacher. In these situations, it is sometimes helpful to evaluate improvement in relation to specific situations and settings. In other words, a child may be making considerable improvement in Mr. Smith's math class but demonstrating no improvement in Mrs. Martin's English class.

One of the problems with the use of rating scales as measures of outcomes pertains to the instruments possessing the necessary sensitivity to detect meaningful change. Changes resulting from effective interventions might not appear in changes in scores on rating scales because the targeted behavior change may have been more specific than can be indicated on most rating scales. For example, interventions may effectively reduce a set of behavior that annoys peers in math class; however, this may produce very little change in the child's overall social functioning or acceptance by other children and adults. As a result, the use of most rating scales as an outcome measure in this situation is probably inappropriate. To measure changes in behavior, it is frequently useful to have repeated measurement of target behavior so as to track change. One of the practical problems with administering rating scales is the timely completion and return of assessments by parents and teachers. It is usually necessary to closely monitor their return, check

for completion (e.g., whether the rater completed the back of the form), and follow-up with raters whose forms are incomplete or late.

Performance Measures

Performance measures assess the ability of a child or an adolescent to exhibit a skill. They are frequently used in academic achievement testing, but they also have applications related to behavioral and emotional problems. Pre- and post-assessments can be administered to assess whether a child had a skill (e.g., social skill, problem-solving skill) prior to the intervention and possesses the skill after the intervention. These assessments tend to be straightforward and based on an appraisal of the child's ability to demonstrate through role-play, written responses, and/or verbal description a set of skills that have been taught directly. These techniques are useful for determining whether a child or an adolescent is ready to begin applying the skills in actual situations because generalization is not likely to occur if the child cannot produce the behavior on cue.

A common misapplication of these measures in evaluating outcomes is an overreliance on them as the end point of the need for interventions. Many socially impaired youth have been successfully taught numerous social skills that have resulted in no change in behavior in other settings. Generalization, or the degree to which learned skills are displayed in novel or untrained situations, is usually the most challenging part of any behavioral intervention, and measures assessing behavior change in the target settings is a necessary outcome measure.

Current Issues

Successfully evaluating outcomes requires giving attention to many issues, including generalization, mediators and moderators, and size of effects. Attention to generalization requires evaluators to focus key outcome measures on the settings in which the problems that led to the initial evaluation occurred. Although measurement of the process of interventions can help to inform treatment decisions, it is possible to understand the ultimate success of interventions only when assessments target the problems in the context in which they have occurred. Moderators and mediators refer to characteristics of the

individuals receiving interventions and characteristics of the intervention process that affect treatment outcomes. For example, interventions provided by different teachers in different classrooms may result in mixed results, not as a result of the intervention but rather as a result of whether or not each teacher implemented the techniques correctly. In this example, it could be said that the implementation of the technique "mediated" the outcomes.

Although data indicating improvements are encouraging, questions will remain as to how much improvement is necessary before it is concluded that meaningful changes have taken place. This issue is at the heart of indexes of effect size and clinical significance. These calculations rely on data from the same assessment techniques as described previously but include analyses to assist in the interpretations of the size of the effects. There are various approaches to this topic, including measures of whether change moved a student from outside of the normal range to within the normal range as well as indexes of the size of change in relation to the standard error of measurement and standard deviation. Unlike many other analyses, these techniques can be applied to data for an individual child or a group of children.

LEGAL AND ETHICAL ISSUES

For states to receive federal funding, there must be strict adherence to the principles of the applicable at the federal level. State laws are often more specific than federal legislation and act to flesh out federal mandates, such as IDEA '97, by defining the means that school professionals use to meet the requirements outlined in these laws. With states operating independently in this fashion, there are often differences among states in terms of how the federal laws are interpreted. One state's procedures for conducting an FBA or a BIP may differ from those of another state. Furthermore, requirements set forth in IDEA that are not addressed by the state may be interpreted and defined at the school district level, sometimes resulting in dramatic differences in assessment and classification procedures between adjacent neighborhoods. Therefore, the onus is placed on schools to train their staff to be able to conduct

behavioral assessment in accordance with the pertinent laws and local procedures.

In terms of behavioral assessment, perhaps the two most relevant areas of legal and ethical concern are informed consent and procedures for ensuring cultural competence in assessment. As with any form of nonacademic assessment in the school setting, legal and ethical issues arise around consent for assessment. IDEA and the American Psychological Association's Ethical Principles of Psychologists specifically require that informed consent be obtained from the parents or guardians of a minor before conducting an initial evaluation. For consent to be "informed," the signor must (a) fully understand the purposes and means of the assessment, (b) provide consent voluntarily, and (c) be legally competent. It is generally assumed that children are incompetent to provide independent consent, although there are situations where exceptions are made. For example, students who are emancipated from their parents may be eligible to give consent on their own behalf, and many states allow independent consent for mental health services under 18 years of age. In general, however, parents or guardians are typically required to consent to any assessment services provided to children.

When obtaining consent, professionals must keep in mind that the school environment creates a unique situation for behavioral assessment because, in contrast to clinical settings, it is not automatically assumed that a child's participation will include such services. The clinical practice of "passive consent," or the assumption of consent based merely on participation, is not applicable in the school setting. In fact, some argue that the school setting is inappropriate for mental health services in general and that the practice of assessing children in a school in domains outside of academia is inappropriate. Therefore, professionals in the school setting have the legal and ethical duty to ensure that parents or guardians understand the purpose and means of any assessment that is conducted. It is strongly recommended that school professionals obtain "active consent," which includes formal documentation of parental consent and meets all of the requirements of consent described

previously. Ultimately, this might mean that students will not receive the services they need due to parental disapproval or apathy. Although this is unfortunate, the need for active consent presents such an ethical, and at times legal, concern that it is suggested that professionals in the schools always seek active consent despite the possible consequences.

Another area of legal and ethical concern is cultural competence. In simplest terms, there are two components of cultural competency. First, professionals must understand and respect the values, beliefs, customs, and traditions of the families they serve. This component of cultural competence involves interacting appropriately with people from diverse cultures, implementing a communication process whereby practitioners interact with people as equals (and not as experts), and demonstrating a genuine commitment to understanding students' strengths and weaknesses in the context of the culture, resources, and environment in which the behaviors occur. As mentioned previously, when a professional uses assessment instruments that are normed on nonrepresentational samples, there is a need to interpret the results with great care.

Cultural competence in this respect may involve the ability to interpret the results within the correct framework. For example, a child of Hispanic background who appears to be significantly "introverted" or to lack "leadership" skills on some rating scales may in fact exhibit behaviors well within ranges that would be expected of his or her family oriented cooperative culture. It is the culturally competent professional who can interpret these results in the correct context.

The second component to cultural competence is an awareness of how professionals' own cultural backgrounds affect their services to the students. In some cases, whole schools can adopt a white middle-class culture as a result of the dominant cultural makeup of the faculty and staff. In these instances, there may be a tendency within the school to place an emphasis on timeliness, parental involvement (to the exclusion of extended family members), and a lack of emotional display. Although this might not be self-evident among the faculty and staff members

of the school, these cultural norms can be at odds with families of diverse cultural backgrounds. For example, school personnel can mistakenly interpret students who arrive late for an assessment, or parents who arrive late to meetings, as uncaring or uninvolved. However, in a family's culture, there might not be the same emphasis on timeliness. Similarly, students who openly weep at unfortunate news or seem to be unmoved by positive feedback may be mistakenly considered to be unstable or unconcerned. Again, these reactions might not fit with the dominant culture of the school or professionals themselves, but they still fall well within the norm for the student's culture. As a result, professionals must continually assess their own cultural influences and be willing to recognize instances where their own cultural values do not match those of the students and their families.

FURTHER READINGS

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.

Dunlap, G., Kern-Dunlap, L., Clarke, S., & Robbins, F. R. (1991). Functional assessment, curricular revision, and severe behavior problems. *Journal of Applied Behavior Analysis*, 24, 387–397.

Ervin, R. A., Radford, P. M., Bertsch, K., Piper, A. L., Ehrhardt, K. E., & Poling, K. (2001). A descriptive analysis and critique of the empirical literature on school-based functional assessment. *School Psychology Review*, 30, 193–210.

Ghuman, H. S., Weist, M. D., & Sarles, R. M. (Eds.). (2002). *Providing mental health services to youth where they are*. New York: Brunner–Routledge. Gresham, F. M., Watson, T. S., & Skinner, C. H. (2001). Functional behavioral assessment: Principles, procedures, and future directions. *School Psychology Review*, 30, 156–172.

Mash, E. J., & Terdal, L. G. (Eds.). (1997). *Assessment of childhood disorders* (3rd ed.). New

York: Guilford. Pellegrini, A. D. (2001). Practitioner review: The role of direct observation in the assessment of young children. *Journal of Child Psychology and Psychiatry*, 42, 861–869.

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

Sterling-Turner, H. E., Robinson, S. L., & Wilczynski, S. M. (2001). Functional assessment of distracting and disruptive behaviors in the school setting. *School Psychology Review*, 30, 211–226.

Weist, M., Evans, S. W., & Lever, N. (Eds.). (2003). *Handbook of school mental health: Advancing practice and research*. New York: Kluwer/Plenum.