

Commentary: School-Based Observations of Children at School: Promise with Prudence

Mark D. Rapport
University of Central Florida

Psychological assessment traditionally entails assessment processes and measures to address questions about children. In school settings, these questions usually center on why children experience difficulty learning, engage in inappropriate classroom deportment, and fail to develop adaptive peer relationships. The current mini-series (Volpe & McConaughy, 2005) provides a focused review of one of these methods (*viz.*, direct observation) and the myriad measures and applications of observational assessment appropriate for school psychology practice. Contributing authors detail direct observational assessment procedures and measures appropriate for a variety of school settings, including classrooms (Eckert, Martens, & DiGennaro, 2005; Volpe, DiPerna, Hintze, & Shapiro, 2005), playgrounds (Leff & Lakin, 2005), and test sessions and clinical interviews (McConaughy, 2005). In addition, Hintze (2005) discusses overarching psychometric issues regarding observational assessment. The purpose of this commentary is to provide a clinical perspective on some of the central issues and common themes of the series in relation to child assessment in the school.

Children who perform adequately and behave appropriately in school are usually not singled out for observation with rare exception (*e.g.*, gifted or special programs, advanced placement classes). Students who are scholastically challenged—and particularly those with co-occurring behavior problems—represent the largest subset that commands attention in

school settings, and behavioral observation techniques are a vital tool in the psychologist's armamentarium for understanding these difficulties. Behavior problems are essentially actions that children undertake toward objects—including social interactions and educational tasks—that are considered undesirable by others, particularly adults. On many occasions, functional assessment of the child's behavior in the classroom or in other school settings such as the playground provides critical information concerning stimuli, events, and setting characteristics that evoke and/or maintain problem behavior. An understanding of these contingent relationships is essential for the successful design of intervention strategies (see Eckert et al., 2005). On other occasions, careful scrutiny of children's test-taking behavior or behavior during special situations such as clinical interviews provides hints for understanding their difficulties (see McConaughy, 2005). Additional experiences can be designed subsequently to isolate the functional nature of these cognitively challenging conditions or incidences of atypical behavior.

Other academic and behavior problems observed at school originate elsewhere, and the most rigorous functional analysis will fail to identify their source or variables contributing to their maintenance. In these cases, school-based interventions will mostly fail or minimally improve children's functioning until a better understanding of their clinical status is achieved and appropriate intervention implemented. This commentary focuses on these

enigmatic cases, highlighting the contribution of observational assessment instruments and procedures in school-based settings that contribute to understanding children's academic difficulties and behavioral-emotional problems as they relate to clinical status. A clinical perspective necessarily invokes consideration and discussion of clinical disorders and diagnostic processes, which may appear tangential to traditional school-based assessment and intervention practices. Practitioners are urged to consider these issues as part of a broader problem-solving approach continuum of clinical and school-based practices that are necessary to understand children and meet their needs.

Clinical Presentation of Behavior Problems in School

Myriad examples of clinical problems exist for which the origins are significantly influenced by genetics, in utero incidents, biological correlates, and life events, ranging from moderately upsetting to catastrophic. When these problems represent more than an isolated incidence of problematic behaviors and occur in combination with other behavioral problems (i.e., evince a strong covariance pattern), they suggest the presence of a symptom-complex or syndrome. A clinical disorder is suspected when this co-varying set of behavior and/or emotional problems significantly impairs the child's ability to function in one or more settings and is associated with a specifiable onset, course, duration, outcome, response to treatment, and familial, psychosocial, and biological correlates.

Differences among the most common clinical disorders of childhood with respect to age of onset, duration, and course of behavioral-emotional problems are highlighted in Table 1. Appreciation of the subtle differences obtained from historical information can be quite revealing and strengthen clinical acumen. The onset of complaints for most clinical disorders of childhood can be quite variable and represents one of the most valid means for distinguishing among childhood disorders. For example, a child with a relatively benign history of behavioral and academic problems whose academic performance is compromised

beginning in the fifth grade is unlikely to have attention-deficit/hyperactivity disorder (ADHD) based on current conceptualizations of the disorder. More likely alternative explanations might be an affective or anxiety disorder, early onset schizophrenia, abrupt environmental change, or prevailing environmental conditions, such as poor instructional strategies.

Recognizing the presence of a clinical disorder is paramount—regardless of one's vernacular for describing the phenomenon—for selecting an effective intervention. Careful design of extant contingencies within a classroom may improve some aspects of a child's daily functioning, but is frequently insufficient to bring about significant and lasting change in children with serious mental health problems (e.g., ADHD, depression, obsessive-compulsive disorder). Empirically based treatments, such as those developed for depression and certain forms of anxiety, pharmacological regimens for ADHD, discrete trial training for autism, and parent effectiveness training for oppositional-defiant disorder may be needed to complement school-based interventions (for a review of evidence-based treatments, see Kazdin & Weisz, 2003). Comprehensive assessment and diagnosis are thus viewed as complementary components of the full spectrum of clinical and school-based practices that provide a broad perspective of children's behavioral and emotional difficulties.

Best Practice Uses of Direct Observation Methodology in School Settings

Numerous advantages and uses of direct observation methodology were outlined by the mini-series contributing authors. Direct observation of classroom, playground, and test-taking or interview behavior can also be used as an initial step or complementary procedure to obtain critical information concerning children's clinical status when the information is integrated with data obtained from other sources. The key is to recognize the potential contributions of direct observations, while appreciating their inherent limitations. Currently, there is no instrument, process, or technique

Table 1
Onset, Course and Duration of Major Clinical Disorders of Childhood

Clinical Disorders	Onset ^{a,b,c}	Course	Duration
Disruptive Behavior Disorders			
Attention-Deficit Hyperactivity Disorder (ADHD) ^b	3.5 ^b	Chronic	Adoles-Lifelong
Conduct Disorder (CD) ^d	<10 ^c <16 ^c	Variable Variable	Adulthood Early Adult
Oppositional Defiant Disorder (ODD)	<8 ^c	Variable	Remits or Antecedent to CD
Pervasive Developmental Disorders			
Asperger's Disorder	3-6 ^c	Chronic	Lifelong
Autistic Disorder ^e	<3 ^b	Chronic	Lifelong
Childhood Disintegrative Disorder ^e	3-4 ^{b or c}	Chronic	Lifelong
Rett's Disorder ^e	1-2 & <4 ^c	Chronic	Lifelong or Fatal
Mood Disorders			
Major Depressive Disorder ^f	5-19 ^{b or c}	Variable	Remits or Variable
Dysthymic Disorder ^g	8.5 ^c	Variable	Remits or Variable
Mania Episode (In context of Bipolar Disorder)	5-14	Variable	Lifelong
Anxiety Disorders			
Acute Stress Disorder	Any Age ⁱ	1 Month	2-Days to 1-Month
Obsessive-Compulsive Disorder ^j	6-15(m) ^c 20-29(f) ^c	Chronic Chronic	Lifelong Lifelong
Posttraumatic Stress Disorder	Acute or Delayed ⁱ	Variable Variable	2-Months to 2-Years
Separation Anxiety Disorder ^m	9-13 ^{b or c}	Variable	2-Yrs-Adoles
Social Phobia ^m	Mid-Teens ^{b or c}	Chronic	Remits by Adulthood
Specific Phobia ^{k,m}	7-12 ^{b or c}	Variable	Remits by Adoles
Other Clinical Disorders			
Tourette's Disorder	7 ^c	Variable	Lifelong
Early Onset Schizophrenia	5-11 ^{b or c}	Variable	Lifelong

Note: ^a age of onset indicates age in years at which symptoms are most frequently first reported in children; ^b acute onset; ^c insidious onset; ^d at risk for antisocial personality disorder and substance abuse disorder as adults; ^e typically associated with an anxiety disorder; ^f frequently associated with an anxiety disorder; ^g frequently associated with externalizing problems; ^h frequently associated with oppositional defiant and conduct disorder; ⁱ onset immediately following a traumatic event; ^j commonly associated with depression, other anxiety disorders, and/or Tourette's; ^k slightly higher rates in females and dependent upon the type of phobia; ^l significantly higher number of males versus females prior to age 10; ^m frequently continuous with adult anxiety disorder; Adoles = adolescence; (f) = females; (m) = males.

that is sufficiently sensitive to enable its use in isolation in assessment of children's functioning. Instead, evidence-based diagnosis is recommended, wherein reliable and valid information is obtained from multiple sources pertaining to a child's past and current functioning, while minimizing subjective clinical judgment (for a review, see Mash & Hunsley, 2005). In support of this recommendation, recent studies have shown weak, if any, association between evidence-based diagnosis and unstructured assessments guided only by clinical judgment. For example, less than chance level agreements were reported between clinician chart review of youths treated in community mental health clinics and diagnosis generated by a structured diagnostic interview (Jensen & Weisz, 2002). Other researchers also reported similar results comparing clinically based diagnosis with evidence-based diagnosis (Lewczyk, Garland, Hurlburt, Gearity, & Hough, 2003).

Direct observation methodology's contribution to evidence-based diagnosis holds clear promise based on the discriminant validity estimates reported for a wide range of instruments in this series. Additional psychometric properties must be established to broaden their application for evidence-based diagnosis as discussed below.

Contribution of Direct Observations to Differential Diagnosis

Commonalities and Nonspecificity of Behavior-Emotional Problems

Some behavior problems are common to nearly all clinical disorders and complicate the differential diagnostic process. That is, their presence is not pathognomic of any particular disorder (i.e., low positive predictive power), but their absence indicates the improbability of a disorder (high negative predictive power). Inattentiveness, or the inability to pay attention in the classroom for an age-appropriate duration of time, is a stellar example. Difficulties with attention are consistently reported in children with ADHD, but are also apparent in children with pervasive developmental disorders, childhood depression, most forms of anxiety

disorder, mental retardation, conduct disturbance, early onset schizophrenia, a variety of adjustment reactions (e.g., bereavement, parental divorce, relocation), and various medical diseases (e.g., seizure disorders). Pharmacological regimens (e.g., antipsychotics), chronic abuse, environmental toxins (e.g., lead paint, second-hand smoke), insufficient sleep, inadequate diet, and excessive computer use (e.g., internet messaging and gaming) and television viewing are also associated with attention problems in children. Four metrics address this overlap: sensitivity, specificity, positive predictive power (PPP), and negative predictive power (NPP).

Sensitivity indicates the proportion of a studied group with a target diagnosis who test positive on a specific measure. Specificity indicates the proportion of the same group without a target diagnosis who test negative on a specific measure. These indices are useful for examining the overall classification accuracy of total and subscale scores derived from the direct observation systems and rating scales, and appear promising based on the strong discriminant validity estimates reported in this series of articles. The metrics most relevant to practitioners, however, are positive (PPP) and negative (NPP) predictive power. PPP, as it applies to direct observations and rating scales, indicates the conditional probability that a child exceeding an established cutoff score or exhibiting particular behaviors while being observed or interviewed meets criteria for a particular diagnosis, such as ADHD (i.e., the ratio of true positive cases to all test positives). NPP, in contrast, indicates the conditional probability that a child who does not exceed an established cutoff score or fails to display particular behaviors will not meet criteria for a particular diagnosis (i.e., the ratio of true negative cases to all test negatives). As an example, excessively low rates of eye contact and spontaneous speech observed during a clinical interview or high rates of aggressive behavior on the playground, may portend the presence of significant and pervasive behavioral or emotional problems that extend beyond the school setting and are worthy of further investigation. High values (e.g., $>.80$) for all four indices are desirable.

Empirical research is needed to establish sensitivity, specificity, PPP, and NPP for the growing number of direct observation instruments. Doing so will broaden their usefulness as evidence-based assessment instruments that can be incorporated into a comprehensive diagnostic evaluation for children suspected of significant mental health problems.

Validity's Uncertainties

Some commonly used interviews, rating scales, and clinic-based instruments estimate validity based on how well they correlate with other rating scales and instruments, or in the case of clinical interviews, with clinician (i.e., sometimes termed, "expert") diagnosis. This approach is inadequate for demonstrating convergent validity when the rating scale uses near identically worded descriptions of behavior problems (e.g., items based on DSM-IV criteria or nearly identical items borrowed from existing instruments). Essential to construct validation is the requirement that scale or interview items be evaluated against independent measures of the construct. The same holds for comparisons between clinical interviews and expert judges' opinions. Clinicians are not particularly reliable at assessing diagnosis, and the validity of a diagnostic instrument must be independent from the diagnostic criteria it assesses—that is, research must distinguish between the validity of a disorder and the criteria used to define it.

Unwarranted confidence may also result from the practice of estimating discriminant validity based exclusively on correlated score patterns with instruments purportedly measuring the same construct. Extant research reveals that even when scale scores are significantly correlated, they may be unrelated to objective measures (for a review, see Rapport, Kofler, & Himmerich, in press). For example, previous studies assessing activity level in children reveal that those receiving higher teacher activity ratings than other children in the same classroom were, in fact, less motorically active according to precision counters used concurrently to measure motor movement—approximately 64% of children rated as "clinically hyperactive" were less active than the

most active child rated as being normal by the teacher (Tryon & Pinto, 1994).

A separate set of problems exists because of the lack of published norms for judging whether children exceed developmentally appropriate levels of behavior. Although norms are available for several rating scales and direct observation instruments, they typically encompass broad age spans (e.g., 6 to 12 years of age) that fail to reflect the rapid developmental changes observed in children. Several types of commonly occurring behavior problems decrease with age (e.g., lying, activity level), some behaviors and a wide range of cognitive abilities show rapid developmental change from year to year (sustained attention, metacognitive skills), and the typography rather than frequency changes for other behaviors as children mature (e.g., aggression). Collectively, these shortcomings highlight the need for multimethod, multisource, multisetting information coupled with historical information and appropriate age norms when evaluating children. Excellent examples of multimethod behavioral assessment approaches that encompass both direct and indirect methods for school settings are available (Mash & Terdal, 2000; Shapiro & Kratochwill, 2000). Direct observation methodology that includes typically developing classroom comparison children bypasses many of the potential shortcomings, and is recommended when normed age ranges span more than a year or two.

Concluding Comments

Information obtained by means of direct observation methodology, whether by observing children's classroom behavior, social interaction pattern on the playground, or demeanor while being interviewed or tested, provides an important snapshot of children's lives. Legitimate questions must be addressed concerning how far we can generalize from these observations to making causal or probability statements about children's behavioral-emotional problems and competencies—particularly those with significant mental health problems. To be most effective, evidence-based diagnostic practices must include information from multiple sources, as well as systematic

direct observation, to provide comprehensive pictures of children's current functioning in school and their needs for school-based interventions.

References

- Eckert, T. L., Martens, B. K., & DiGennaro, F. D. (2005). Increasing the accuracy of functional assessment methods: Describing antecedent-behavior-consequence relations using conditional probabilities. *School Psychology Review, 34*, 520-528.
- Hintze, J. M. (2005). Psychometrics of direct observation. *School Psychology Review, 34*, 507-519.
- Jensen, A. L., & Weisz, J. R. (2002). Assessing match and mismatch between practitioner generated and standardized interview-generated diagnoses for clinic-referred children and adolescents. *Journal of Consulting and Clinical Psychology, 70*, 158-168.
- Kazdin, A. E., & Weisz, J. R. (Eds.). (2003). *Evidence-based psychotherapies for children and adolescents*. New York: Guilford.
- Leff, S. S., & Lakin, R. (2005). Determining best practice guidelines for playground-based observational research: Implications for practitioners and researchers. *School Psychology Review, 34*, 474-488.
- Lewczyk, C. M., Garland, A. F., Hurlburt, M. S., Gearity, J., & Hough, R. L. (2003). Comparing DISC-IV and clinician diagnoses among youth receiving public mental health services. *Journal of the American Academy of Child and Adolescent Psychiatry, 42*, 349-356.
- Mash, E. J., & Hunsley, J. (2005). Evidence-based assessment of child and adolescent disorders: Issues and challenges. *Journal of Clinical Child and Adolescent Psychology, 34*, 362-379.
- Mash, E. J., & Terdal, L.G. (Eds.). (2000). *Assessment of childhood disorders* (3rd ed.). New York: Guilford Press.
- McConaughy, S. H. (2005). Direct observational assessment during test sessions and child clinical interviews. *School Psychology Review, 34*, 489-506.
- Rapport, M. D., Kofler, M. J., & Himmerich, C. (in press). Activity level in children. In M. Hersen (Ed.), *Clinician's handbook of child behavioral assessment*. New York: Elsevier Inc.
- Shapiro, E. S., & Kratochwill, T. R. (Eds.). (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.
- Tryon, W. W., & Pinto, L. P. (1994). Comparing activity measurements and ratings. *Behavior Modification, 18*, 251-261.
- Volpe, R. J., DiPerna, J. C., Hintze, J. M., & Shapiro, E. S. (2005). Observing students in classroom settings: A review of seven coding schemes. *School Psychology Review, 34*, 454-473.
- Volpe, R. J., & McConaughy, S. H. (2005). Systematic direct observational assessment of student behavior: Its use and interpretation in multiple settings: An introduction to the miniseries. *School Psychology Review, 34*, 451-453.

Mark D. Rapport received his PhD in Clinical Psychology and School/Community Psychology (joint major) from Florida State University in 1981. He is currently a Professor of Clinical Child Psychology, and Director of the Children's Learning Clinic-IV at the University of Central Florida in Orlando. His primary research interests center on understanding underlying core deficits such as working memory in children with ADHD through conceptual model validation.

Copyright of School Psychology Review is the property of National Association of School Psychologists and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.