Reliability and Validity of the Direct Observation Form of the Child Behavior Checklist

Michael L. Reed and Craig Edelbrock
University of Pittsburgh School of Medicine

This article reports reliability and validity data for the Direct Observation Form (DOF) of the Child Behavior Checklist. Observational data were collected on two samples of boys aged 6-11 in classroom settings. Interobserver agreement was high: r = .92 for behavior problem score and r = .83 for on-task score. Generalizability, as measured by the one-way intraclass correlation, was .86 and .71 for behavior problem score and on-task score, respectively. In terms of validity, DOF scores correlated significantly and in the expected directions with teacher-reported problem behavior, school performance, and adaptive functioning. In addition, boys who had been referred by their teachers due to problem behavior obtained significantly higher behavior problem scores and significantly lower on-task scores than a matched sample of normal boys observed in the same classrooms.

Observational assessment is useful for identifying children having emotional and/or behavioral disorders, for making clinical and educational decisions regarding disturbed children, and for measuring changes in

Manuscript received in final form March 16, 1983.

1 A portion of this research was included in a dissertation submitted by the first author to the Graduate School of Oregon State University in partial fulfillment of requirements for the Ph.D. degree. This research was supported in part by the Eric Englund Scholarship to the first author, by an NIMH Research Scientist Development Award (MH00403) to the second author, and by NIMH grant number MH37372. The generous support of the William T. Grant Foundation is gratefully acknowledged. The authors are indebted to the Ralston Public Schools and Corvallis Public Schools for their cooperation and to Dr. Alan Sugawara for his help in this research. The authors wish to thank Dr. Thomas Achenbach for his advice regarding this research and his comments on earlier versions of this manuscript.

2 Address all correspondence to Michael L. Reed, WPIC, 3811 O'Hara Street, Pittsburgh, Pennsylvania 15213.
behavior over time. Observational measures are also useful for evaluating interventions designed to reduce problem behavior and foster positive adjustment, and they have the advantage that observers can be kept blind to the treatment status of target children. Several child behavior coding systems have been developed, some of which have adequate reliability and validity. Although excellent in many ways, these instruments are of limited value for certain purposes because they involve special data-collection equipment or complex coding procedures (e.g., Ray, Shaw, & Patterson, 1968; Patterson & Reid, 1973), require lengthy observer training (e.g., Wahler, House, & Stambaugh, 1976), or are focused upon narrowly defined syndromes such as acting out (Werry & Quay, 1969), hyperactivity (Abikoff, Gittelman-Klein, & Klein, 1977), anxiety (Glennon & Weisz, 1978), or autism (Freeman, Ritvo, Guthrie, Schroth, & Ball, 1978).

The Direct Observation Form (DOF) of the Child Behavior Checklist (CBCL) was designed to fill the need for a simple and efficient observational assessment system that does not require special data-collection equipment, clinically sophisticated observers, or lengthy observer training. The DOF can be used in applied settings, such as the classroom, and it encompasses a broad range of problem behaviors. In addition, the DOF parallels the parent and teacher versions of the CBCL (Achenbach, 1978; Achenbach & Edelbrock, 1979; Edelbrock & Achenbach, Note 1) in terms of item selection and wording. Of 96 items on the DOF, 86 overlap with items on the teacher CBCL, whereas 73 overlap with items on the CBCL completed by parents. This makes it possible to integrate observational data more readily with information from these other informants.

This study reports preliminary reliability and validity data for the DOF. Classroom data from two samples of children were used to evaluate interobserver agreement, determine the generalizability of observational ratings, and determine the convergent and discriminant validity of DOF scores.

METHOD

Instrumentation

Direct Observation Form. The DOF comprises 96 behavior problem items and a measure of on-task behavior. Assessments of problem behavior are obtained via the following procedure: The child is observed for 10 minutes. During the observational period the observer writes a narrative description of the child's behavior. Although the resulting narrative is not directly scored, this procedure maintains attention on the target child and enables the observer to note the occurrence, duration, and intensity of any