VALIDITY OF SELF- AND PARENT-REPORT SCALES IN SCREENING STUDENTS FOR BEHAVIORAL AND EMOTIONAL PROBLEMS IN ELEMENTARY SCHOOL

NANCY HOWELLS WROBEL
University of Michigan-Dearborn

DAVID LACHAR
University of Texas Medical School at Houston

The comparative validity of the child-report Classroom Screening Scale (CLASS) from the Personality Inventory for Youth (PIY), and the parent-report Adjustment Scale (ADJ) from the Personality Inventory for Children (PIC) was examined. The subjects were 111 children from fourth- through sixth-grade regular education classrooms. Children were classified as positive or negative for risk status on each screening scale using a cutoff of 60 T. Parent description identified a higher proportion of children as at risk. Those identified as “at risk” by either parent or child report were more likely than others to have clinically elevated scores on both the PIC and PIY clinical scales. High scores on CLASS were more likely to have clinically significant elevations than were high scorers on ADJ. Parent report appeared to be sensitive to overt behavioral problems, while self-report was sensitive to mood disturbances and withdrawal. © 1998 John Wiley & Sons, Inc.

The need for early identification and intervention in the treatment of childhood psychopathology coupled with limited resources for clinical evaluation in schools require efficient screening instruments. School psychologists are often faced with the task of evaluating and treating children who are referred by school personnel due to behavioral and emotional problems. Nodrick and Li (1992) note that these referrals for psychological services are often presented in very general terms and lack the specificity of detail even necessary to determine the need for a formal evaluation. Increasingly, school and clinical psychologists are turning to instruments such as behavioral rating scales for the initial assessment of children’s psychological problems (Hutton, Dubes, & Muir, 1992). This increase parallels the increased interest in applying objective methods to assess behavioral and emotional adjustment of children and adolescents over the past two decades (Lachar, 1993).

MULTIPLE SOURCES OF ASSESSMENT INFORMATION

Teacher Report

A primary source of descriptive information has been teacher reports. Given that teachers are the most frequent referral source in a school setting, they are also likely sources of information about the severity and specific nature of children’s problems. Ratings by teachers typically focus on elementary school children who require the evaluation of concentration, activity level, or behavioral control (Lachar, 1993).

Teacher reports may vary depending upon the context in which the child is observed and the utility of teacher reports may also vary with the type of behavioral problem. A survey of mental health professionals revealed that teachers were viewed as useful informants of child hyperactivity and inattentiveness but as less useful informants than mothers and children with regard to internalizing behaviors (Loeber, Green, & Lahey, 1990).

Parent Report

Research using both the Teacher Report Form and the Child Behavior Checklist (Achenbach, 1983) suggest that while teachers provide some useful information, parent observation is perhaps the more useful in screening for some types of problems (Bird, Gould, Rubio–Stipec, Staghezza, & Cani-
Parents are considered important sources of information because they are most likely to observe the child over long periods of time and in many contexts (McConaughy, 1993). When surveyed regarding parents as informants, mental health professionals tended to view parents as optimal informants with regard to oppositional behaviors and somewhat useful with regard to internalizing symptoms (Loeber, Green, & Lahey, 1990). Parents appear to be generally good at reporting behavior and conduct problems (Dollinger, 1992; Edelbrock, Costello, Dulcan, Conover, & Kala, 1986; Herjanic & Reich, 1982; Verhulst & van der Ende, 1991). Parents appear to be less useful sources of information when depressive symptoms are of concern (Apter, Orvaschel, Laseg, Moses, & Tyano, 1989; Angold, Weissman, John, Merikangas, Prusoff, Wickramaratne, Gammon, & Warner, 1987; Kashani, Orvaschel, Burk, & Reid, 1985). The tendency is for parents to underreport depression; however, when they do report such symptoms their reports tend to be high in specificity (Angold et al., 1987) and are related to overt behavioral measures of depression (Kazdin, Esveldt-Dawson, Sherick, & Colbus, 1985).

Child Report

A third source of diagnostic information is the child. Child report is not usually considered useful in screening for psychological problems, because children are typically referred by adults who see their behavior as “disturbed or disturbing” (Algozzine, 1977) rather than self-referred. Use of self-report for screening school children may be advantageous for two reasons. Unlike in clinical settings, where parents would typically accompany the child, the parent or guardian may not be readily available to school personnel for completion of a comprehensive questionnaire. Secondly, there are certain problems about which children may be more knowledgeable or more willing to report about themselves than are parents. Although children may be less accurate reporters of attentional problems or hyperactivity (Loeber, Green, Lahey, & Stouthamer-Loeber, 1990), they have been demonstrated to be the preferred source for some symptoms. Stavrakaki and associates (Stavrakaki, Williams, Walker, Roberts, & Kotsopoulos, 1991) demonstrated that self-report measures including the Revised Children’s Manifest Anxiety Scale (Reynolds & Richmond, 1985) and the Children’s Depression Inventory (Kovacs, 1985) proved to be good screening tools for prepubertal school children. Similarly, Hodges (1990) found prepubertal psychiatric inpatients to be useful informants for screening purposes. Research on the Social Adjustment Inventory for Children and Adolescents (SAICA; John, Gammon, Prusoff, & Warner, 1987) indicates that while mothers and children generally tend to agree on reports of social adjustment, children are more likely to identify problems in this area. In addition, the child may be the best informant of problems about which the parent is unaware, such as delinquent behaviors or internal symptoms, such as disturbed thought processes.

The capacity and willingness of children to provide self-report may vary somewhat with the child’s age. Generally children appear to be more capable cognitively of assessing their own behavior as they increase in age (Edelbrock et al., 1986), although the capacity to conceal behaviors also increases with age. Thus, in a study by Achenbach and associates (Achenbach, McConaughy, & Howell, 1987), correlations between parent and child report were higher for 6–11-year-olds than for adolescents. Given the perhaps untapped capacity and willingness for older elementary school children to provide information, these findings highlight the need for instruments which can be utilized to allow younger children to report their behaviors and symptoms for screening purposes. The present study examines the usefulness of such an instrument, namely the Personality Inventory for Youth (PIY). The PIY has recently been developed as a self-report version of the Personality Inventory for Children (PIC; Wirt, Lachar, Klinedinst, & Seat, 1984).

This study addresses the validity of PIY relative to the PIC Adjustment scale (ADJ). ADJ is a parent-report screening scale empirically derived to identify children in need of psychological evaluation and as a general indicator of adjustment problems. The study examines whether the child
self-report screening instrument (PIY CLASS scale) provides information which is equally useful in the early screening process. Both CLASS and ADJ scores will be examined for their ability to predict elevations on the clinical scales of the PIY and PIC. This allows an opportunity to observe which behaviors prompt a child or parent to identify the child as in need of intervention, based on the screening scales. It is predicted that children will be capable of identifying themselves as in need of further assessment; however, they may be sensitive to different problems than parents. In particular, it is expected that children identified as ‘at risk’ based on either CLASS or ADJ screening scales will have elevated profiles on a greater number of scales than those without elevated screening scales; however, those identified as ‘at risk’ by adult informants will have a greater proportion of elevated scores on clinical scales representing more externalizing symptoms (e.g., Delinquency and Hyperactivity). Children who are self-identified as ‘at risk’ will have a greater proportion of elevations on scales representing more internalizing behaviors (e.g., Depression and Withdrawal).

Use of the clinical scales from the PIC and PIY as criterion measures was deemed useful for two reasons. First, the correspondence between the PIY and PIC scales allows both parent and child to respond to similar items. Thus, where opinions diverge (for instance, the child reports one type of problem and the parent another) we may be able to view these distinctions as possibly due to different perspectives rather than due to the use of extremely disparate instruments. Secondly, although CLASS items are drawn from the same pool of items as the PIY clinical scales, the items represent a wide range of symptoms; therefore, overlap with any specific scale is fairly minimal. The greatest overlap occurs with the PIY Delinquency Scale, contributing 12 of the 32 items. The remaining clinical PIY scales each share between two and four items with CLASS.

**METHODS**

**Subjects**

The subjects were 111 children in regular fourth- through sixth-grade classrooms at three midwestern public elementary schools in middle class and upper middle class suburbs. They ranged in age from 9 to 13 with a mean age of 10.35. Gender was equally represented, with 56 males and 55 females. The children were predominantly white (90.1%). Children were excluded from the study if they did not complete a valid PIY. PIY profiles were considered invalid if the Inconsistency score was greater than 70T (two such cases were eliminated) or Dissimulation was greater than 80T (four cases were eliminated).

**Measures**

*The Revised Personality Inventory for Children.* The PIC is an objective multidimensional test of child and adolescent behavior, emotional, and cognitive status completed by a parent or other informant. Lie (L), Frequency (F), and Defensiveness (DEF) scales parallel the L, F, and K scales of the MMPI. Adjustment (ADJ) scores represent the general level of psychological disturbance. Clinical scales include: Achievement (ACH), Intellectual Screening (IS), Development (DVL), Delinquency (DLQ), Hyperactivity (HPR), Somatic Concern (SOM), Depression (D), Withdrawal (WDL), Anxiety (ANX), Psychosis (PSY), Social Skills (SSK), and Family Relations (FAM). PIC scales IS, SOM, D, WDL, ANX, and SSK are considered clinically elevated if T ≥ 70. PIC scales DLQ and PSY are clinically elevated if T ≥ 80. The remainder are elevated if T ≥ 60. Studies of the reliability of PIC results yield 2-week test–retest coefficients averaging .86 (Wirt et al., 1984). The PIC clinical scales were developed using one of two methods: empirical item selection based on the analysis of criterion samples or other response frequency criteria, or content scale methods that incorporate consensual nomination of items to reflect specified dimensions, followed by statistical refinement aimed at enhancing internal consistency. Establishment of the predictive validity and in-
interpretation of these scales was conducted by Lachar and Gdowski (1979) utilizing external criteria including parent, teacher, and clinician’s ratings. A recent review of test performance and application may be found in Lachar and Kline (1994).

As noted, the PIC ADJ scale represents a parent-report screening scale intended to identify children with emotional and other psychological problems. Items were selected by contrasting the protocols of normal boys with those of boys identified as having emotional disturbance. This 54-item scale has demonstrated stability as well as predictive validity (Wirt et al., 1984; Lachar, 1982).

The criterion utilized in the validation of ADJ included independent correlate forms completed by parents, teachers, and clinicians. Parents’ correlates of ADJ reflected a broad scope of concerns which might motivate a parent to seek psychological intervention. Increasing elevations reflected a greater variety and frequency of problems. Correlates provided by teachers suggest that ADJ reflects poor school performance as well as poor adjustment to the school environment. Reports obtained from clinicians indicated that ADJ reflects a variety of “externalizing, acting-out” behaviors, including hostility, dyscontrol, and difficulties accepting limits (Lachar & Gdowski, 1979).

**Personality Inventory for Youth.** This is a self-report inventory derived in large part from the first 280 items of the PIC. The majority (61%) of items were transformed into first person statements with little change. One third of the PIC-R items required more substantial revision. The remaining items were discarded because they reflected early developmental milestones of which the child may be unaware or behavior below the developmental level of the intended population. It is written at a grade level of 3.4 (Fry, 1968; 1972) with norms including regular education students in fourth through twelfth grade. The PIY consists of the Classroom Screening Scale as well as nine nonoverlapping clinical scales and 24 complimentary subscales which assess cognitive as well as behavioral and emotional difficulties. These clinical scales include Cognitive Impairment (COG), Impulsivity and Distractability (ADH), Delinquency (DLQ), Family Dysfunction (FAM), Reality Distortion (RLT), Somatic Concern (SOM), Psychological Discomfort (DIS), Social Withdrawal (WDL), and Social Skill Deficits (SSK). All PIY scales are considered elevated if \( T \geq 60 \). Test-retest reliability and coefficient alphas conducted on these scales were mainly in the .80s. In addition, four validity scales have been derived to detect the presence of defensive, careless, exaggerated, or noncompliant response (Lachar & Gruber, 1995).

The Personality Inventory for Youth (PIY) Classroom Screening Scale (CLASS; Lachar & Gruber, 1995) is a self-report scale designed as a first stage screen to identify emotional and behavioral disturbance among regular education students. It is envisioned as the possible first gate in a multiple source approach to screening and assessment. CLASS is a 32-item empirically derived scale. Items come from all nine clinical scales; however, the greatest overlap is with the Delinquency scale. Scale construction for CLASS not only involved contrasting clinically referred with regular education groups, but also involved item analysis within the regular education sample. Examination of the internal consistency and temporal stability of CLASS supports its reliability. Factor analytic study of the scale suggested three strong elements including internalizing psychopathology, externalizing psychopathology, and interpersonal adjustment. Previous study of the concurrent validity of CLASS with clinically referred samples revealed modest correlations with parent reports of noncompliance and problematic adjustment (Lachar & Gruber, 1995). CLASS items are included within the first 80 items of the PIY administration booklet, allowing a brief classroom evaluation that can be completed in under 15 minutes (Lachar & Gruber, 1995).

**Procedures**

Elementary school students were administered a prepublication version of the PIY in their classrooms, assisted by an audiotape. Parents completed the 280-item version of the PIC. Each PIC pro-
tocol was classified by the PIC ADJ elevation and each PIY protocol was classified by the PIY CLASS T-score. Assigned risk status for both scales was “low” for those having T-scores less than 60 and “high” for T-scores of 60 or more. Thus a child scoring over 60 on the PIY CLASS scale was viewed as ‘at risk’ based on the child’s self-report, and those over 60 on the PIC ADJ scale were viewed as ‘at risk’ based on parent-report. The 60 cut-off score for CLASS was chosen because the scale effectively separated clinically referred from regular education samples at this level (Lachar & Gruber, 1995). Similarly the 60 cutoff for the PIC ADJ scale was empirically derived through optimal separation of normal versus clinically referred samples (Wirt et al., 1984).

Data Analysis

Mean PIC and PIY profiles were generated for CLASS- and ADJ-based risk groups. Two variables were created by totaling the number of elevated clinical scale scores obtained for each child on the PIY and the PIC. T tests were utilized to compare the total number of elevated PIC and PIY clinical scales for those scoring low versus high on the CLASS and ADJ scales. The proportion of subjects with elevated scores on each PIY and PIC clinical scale was examined within CLASS- and ADJ-assigned groups (low versus high). The χ² statistic, with Yates Correction, was applied to determine whether significantly different proportions of elevated scores (Lachar & Gdowski, 1979; Lachar & Gruber, 1995) were obtained between those scoring high versus low on CLASS and ADJ.

Results

Of the 111 subjects, 6% of subjects were identified as requiring further evaluation based on child-report and 18% were identified based on parent-report. Of those children identified by child-report, 71% were also identified as at risk based on parent-report. The mean CLASS score of those scoring low on CLASS was 42.32 (SD = 4.93), while the mean for high scorers was 65.14 (SD = 5.76). Low scorers on the PIC ADJ scale had a mean ADJ score of 46.32 (SD = 5.16) while high scorers had a mean ADJ score of 72.20 (SD = 11.74).

Figures 1 and 2 depict the mean PIY and PIC clinical scale profiles for children identified as either high or low on the CLASS scales. The means for high CLASS (Figure 1) scorers fell within the clinical range for four PIY scales Impulsivity/Distractability, Psychological Discomfort, Social Withdrawal, and Social Skills Deficits). For those scoring high on CLASS, three PIC scale mean T scores fell within the clinical range (Depression, Family Relations, and Social Skills). The scale mean T scores for the PIY clinical scales of the high ADJ sample (Figure 2) did not fall within the clinical range. The PIC scale means for those scoring high on ADJ were only clinically elevated for the Family Relations scale.

Results of the analysis of the number of elevated clinical scales for those identified as either high or low on the screening scales yielded significant differences for both the ADJ and CLASS scales when compared on PIY and PIC clinical scale elevations. High CLASS scorers had a mean of 3.86 (SD = 2.34) elevated scales on the PIC, while low scorers on CLASS had a mean of 0.63 (SD = 1.22) elevated scales, with a t value of −6.35 (df = 109) and p ≤ .01. On the PIC clinical scales, High CLASS scorers had an average of 4.56 (SD = 1.27) elevations, while low scorers had an average of 0.36 elevations (SD = 0.72). This resulted in a t value of −8.67 (df = 109), and a p ≤ .01. High ADJ scorers had an average of 3.05 (SD = 2.28) elevated PIC clinical scales, versus an average of 0.34 (SD = 0.62) for the low scorers on ADJ (t value = −9.92 (df = 109), p ≤ .01). The high scorers on ADJ also had a higher number of elevated PIY scales (M = 1.80, SD = 2.07) than the low scorers (M = 0.36, SD = 0.85), also yielding a significant difference at the p ≤ .01 level (t value = −5.02).

Table 1 presents the proportion of students with clinically elevated scores on each of the PIY and PIC scales as related to CLASS and ADJ risk classification. Chi square analyses revealed that
Figure 1. Mean PIY (above) and PIC (below) scores by CLASS elevations.
Validity of Self- and Parent-Report

Figure 2. Mean PIY (above) and PIC (below) scores by ADJ elevation.
for 8 of the 12 PIC scales a higher proportion of those categorized as high on CLASS had clinical elevations (Achievement, Depression, Family Relations, Delinquency, Withdrawal, Anxiety, Psychosis, and Social Skills). High scorers on CLASS obtained a higher proportion of clinical elevations on all PIY scales except Family Dysfunction with proportions exceeding 50% for PIY scales Delinquency, Social Withdrawal, Reality Distortion, Impulsivity/Distractability, and Social Skill Deficits. High scorers on ADJ obtained significantly more clinical elevations on all PIC clinical scales except Somatic Concern. A significantly higher proportion of high ADJ protocols obtained clinical elevations on six PIY scales (Cognitive Impairment, Somatic Concern, Delinquency, Reality Distortion, Impulsivity/Distractability, and Social Skill Deficits).

**Table 1**

<table>
<thead>
<tr>
<th>PIC Scale</th>
<th>CLASS &lt; 60</th>
<th>CLASS ≥ 60</th>
<th>ADJ &lt; 60</th>
<th>ADJ ≥ 60</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACH</td>
<td>5.0</td>
<td>42.9***</td>
<td>0.0</td>
<td>40.0***</td>
</tr>
<tr>
<td>IS</td>
<td>5.9</td>
<td>0.0</td>
<td>3.3</td>
<td>15.0*</td>
</tr>
<tr>
<td>DVL</td>
<td>3.9</td>
<td>14.3</td>
<td>0.0</td>
<td>25.0***</td>
</tr>
<tr>
<td>SOM</td>
<td>3.0</td>
<td>0.0</td>
<td>2.2</td>
<td>5.0</td>
</tr>
<tr>
<td>D</td>
<td>4.9</td>
<td>57.2***</td>
<td>1.1</td>
<td>15.0***</td>
</tr>
<tr>
<td>FAM</td>
<td>12.6</td>
<td>42.9***</td>
<td>8.8</td>
<td>45.0***</td>
</tr>
<tr>
<td>DLQ</td>
<td>0.0</td>
<td>14.3***</td>
<td>0.0</td>
<td>5.0*</td>
</tr>
<tr>
<td>WDL</td>
<td>2.0</td>
<td>28.6***</td>
<td>1.1</td>
<td>15.0**</td>
</tr>
<tr>
<td>ANX</td>
<td>6.8</td>
<td>28.6*</td>
<td>3.3</td>
<td>30.0***</td>
</tr>
<tr>
<td>PSY</td>
<td>1.0</td>
<td>42.9***</td>
<td>1.1</td>
<td>15.0**</td>
</tr>
<tr>
<td>HPR</td>
<td>15.4</td>
<td>28.6</td>
<td>11.0</td>
<td>40.0***</td>
</tr>
<tr>
<td>SSK</td>
<td>3.0</td>
<td>71.5***</td>
<td>1.1</td>
<td>35.0**</td>
</tr>
</tbody>
</table>

Note. All PIY clinical scales are considered elevated if $T \geq 60$. PIC scales IS, SOM, D, WDL, ANX, and SSK are elevated if $T \geq 70$. PIC scales DLQ and PSY are interpreted as clinically elevated if $T \geq 80$. The remainder are elevated if $T \geq 60$.

*p ≤ .05. **p ≤ .01. ***p ≤ .001.

**Discussion**

This study examined the ability of a child- and a parent-report screening scale to detect narrowly defined problems in adjustment as independently reported by elementary school students and their parents. Despite possible limitations, including the size of the sample and the use of participants from one geographic location, there seem to be some findings of clinical interest. Overall, children were less likely to identify themselves as at risk on the PIY screening scale; however, those who were self-identified by CLASS were apt to demonstrate more elevated scores for all self-report clinical problem scales.

Children who were identified by the parent-report PIC ADJ as being in need of evaluation tended to be the children who more often described themselves on the PIY clinical scales as having problems including cognitive dysfunction, behavioral problems (such as delinquency and impulsivity), a poor self-concept (Reality Distortion), and social skill deficits. This is based on the higher proportion of clinical elevations on these scales among high ADJ scorers. Parents tended not to identify
children who described themselves as having more behaviorally subtle problems, such as mood disturbances and withdrawal.

When children were self-identified via the PIY CLASS scale as at risk, they were described by parents on the PIC clinical scales as having problems across a variety of areas including academic problems, social problems, depression, and anxiety, and as evidencing emotional lability. In particular, those subjects identified through high PIY CLASS self-reports were more often the students for whom parents ascribed problematic depression or social skill deficits on the PIC. These are symptoms which may be overlooked if one merely relies upon parent report for screening purposes. For instance, previous studies suggest that children’s self-reports are more apt to include depressive symptoms (Stavakaki et al., 1991) and problems in social adjustment (John et al., 1987), while parents are more likely to report more overt behavioral symptoms (Dollinger, 1992; Edelbrock et al., 1986; Verhulst & van der Ende, 1991).

When children identified themselves as high risk via the CLASS scale, parents were not particularly likely to describe them as having problems in cognitive skills, somatic concerns, or inattention and/or hyperactivity based on the PIC scale scores. These results are consistent with the scale pattern obtained by those children who were identified by the parent-report screening scale (ADJ), particularly for somatic concerns. Considering the discrepancy between parent- and student-report of somatic concerns, it is possible that parents are not as aware of such internalizing symptoms compared to problems or concerns manifest by disruptive behavior. This conclusion is supported by the low base rate of parent identified somatic concerns in all risk-defined groups (See Table 3). In contrast, self-report resulted in more elevations on this dimension. Failure of CLASS (and to some extent ADJ) to identify children with intellectual deficits or hyperactivity may in part reflect the nature of this study sample. Children with significant problems in these areas are often excluded from the regular classroom setting by fourth through sixth grade in a suburban middle class school and thus a wide range of such behaviors would not be represented here. Also children may be less likely than parents to report symptoms that are predominantly seen as problematic by adults, such as impulsivity and distractibility. Lastly, previous research (Lachar & Gruber, 1993) suggests that CLASS may be less representative of intellectual deficits or academic deficiencies.

There was also some discrepancy demonstrated between self- and parent-report of significant family problems: When the parent screening scale (ADJ) was elevated, greater parent report of family dysfunction was obtained on the PIC clinical scales. However, these same children were not more likely themselves to report family dysfunction on the PIY self-report. The limited child report of significant family dysfunction in our study may relate to the observation that Lachar and Gruber (1995) found that younger children (below 11 years of age) in the PIY normative sample report fewer difficulties on this scale compared to older children. It is possible that younger children are either not as aware of family problems such as marital conflict, or are more reluctant to report these matters.

Conclusions

These results may support the use of CLASS in evaluating elementary children for emotional and behavioral problems. For those children for whom self-report identified increased risk of poor adjustment, parents described problems in a variety of areas. Most notable in terms of clinical significance were descriptions of depression and problems with social relations. The parent-report screening scale was less sensitive to child depression as well as to social withdrawal. The potential for CLASS to identify children with more behaviorally subtle problems suggests that the PIY may be useful in school settings, particularly when parent-derived ratings are not readily available. It appears that the self-report of elementary school children may compliment the descriptions provided
by their parents, because parents and children appear to provide different but equally useful information. Thus child self-report may be a useful addition to a multisource approach to screening and assessment.

REFERENCES


