Schedule for Affective Disorders and Schizophrenia for School-Age Children—Present and Lifetime Version (K-SADS-PL): Initial Reliability and Validity Data

JOAN KAUFMAN, PH.D., BORIS BIRMAHER, M.D., DAVID BRENT, M.D., UMA RAO, M.D., CYNTHIA FLYNN, M.A., PAULA MORECI, M.S.W., DOUGLAS WILLIAMSON, M.A., AND NEAL RYAN, M.D.

ABSTRACT

Objective: To describe the psychometric properties of the Schedule for Affective Disorders and Schizophrenia for School-Age Children—Present and Lifetime version (K-SADS-PL) interview, which surveys additional disorders not assessed in prior K-SADS, contains improved probes and anchor points, includes diagnosis-specific impairment ratings, generates DSM-III-R and DSM-IV diagnoses, and divides symptoms surveyed into a screening interview and five diagnostic supplements. Method: Subjects were 55 psychiatric outpatients and 11 normal controls (aged 7 through 17 years). Both parents and children were used as informants. Concurrent validity of the screen criteria and the K-SADS-PL diagnoses was assessed against standard self-report scales. Interrater (n = 15) and test-retest (n = 20) reliability data were also collected (mean retest interval: 18 days; range: 2 to 38 days). Results: Rating scale data support the concurrent validity of screens and K-SADS-PL diagnoses. Interrater agreement in scoring screens and diagnoses was high (range: 93% to 100%). Test-retest reliability coefficients were in the excellent range for present and/or lifetime diagnoses of major depression, any bipolar, generalized anxiety, conduct, and oppositional defiant disorder (.77 to 1.00) and in the good range for present diagnoses of posttraumatic stress disorder and attention-deficit hyperactivity disorder (.63 to .67). Conclusion: Results suggest the K-SADS-PL generates reliable and valid child psychiatric diagnoses.


The Schedule for Affective Disorders and Schizophrenia for School-Age Children—Present and Lifetime version (K-SADS-PL) was adapted from the Present Episode version of the K-SADS (K-SADS-P) (Chambers et al., 1985). The K-SADS has been used in numerous clinical, naturalistic follow-up, treatment, psychobiological, family-generic, and epidemiological studies of affective and other child psychiatric disorders. This article describes the K-SADS-PL interview and data examining its psychometric properties.

Several limitations of the K-SADS-P prompted its revision, including (1) failure to obtain lifetime psychiatric history information; (2) omission of a number of important child psychiatric diagnoses (e.g., attention-deficit hyperactivity disorder [ADHD], posttraumatic stress disorder [PTSD], tic disorders); (3) problems with several probes and scoring criteria used to elicit and rate symptoms; and (4) absence of diagnosis-specific impairment ratings to facilitate "caseness" determination (Schwab-Stone et al., 1996). Impetus to revise the K-SADS was also fostered by release of the DSM-IV. In addition, our extensive experience with the instrument highlighted the value of developing a screening interview to facilitate differential diagnoses and expedite administration with patients and normal controls.
There are currently a number of K-SADS versions in circulation. Characteristics of these instruments are outlined in Table 1. There are many similarities among the instruments. All are semistructured integrated parent-child interviews with data from parents and children recorded on a common answer sheet by a single interviewer. This allows for comparison of responses from both informants and prompt querying of discrepancies. Diagnoses are then derived by synthesizing the parent and child data. Consistent with prior empirical work (Herjanic and Reich, 1982), greater weight is typically given to parents' reports of observable behavior and children's reports of subjective experiences. Ultimately, however, it is up to the interviewer to use his or her best clinical judgment when integrating the data.

Each of the K-SADS versions also provides detailed probes for eliciting information about symptoms. The anchor points for scoring items on the K-SADS-L (Lifetime version) and K-SADS-IVR (fourth version, revised) are very similar to those for the original K-SADS-P, with most items rated on a 0- to 6-point scale and the remainder of items rated on a 0- to 4-point scale. In contrast, the majority of the scales of the K-SADS-PL were simplified to 0- to 3-point ratings, as some of the original K-SADS-P rating scales were unreliable in scoring severity of current symptomatology, and all were difficult to apply in rating past episodes of disorder because of their length. The K-SADS-E (Epidemiologic version) provides an intermediate approach to scoring, with current symptomatology rated on a 0- to 4-point scale and past symptomatology rated on a 0- to 2-point scale. In addition to rating individual symptoms, each of the instruments also provides ratings of impairment. The K-SADS-PL, however, is the only instrument that provides global and diagnosis-specific impairment ratings.

Both the K-SADS-PL and K-SADS-E contain "skip-out" criteria for entry into each diagnostic area, alleviating the need to inquire about all symptoms as in the K-SADS-P, K-SADS-L, and K-SADS-IVR interviews. In the K-SADS-E, if a child screens positive for a given diagnosis, he or she is immediately queried about the remaining symptoms associated with that diagnosis. In the K-SADS-PL, all screen questions are surveyed first, then supplemental questions for diagnoses the child screened positively for are administered. As described further in the "Method" section, the Screen Interview is designed to provide a diagnostic overview of lifetime psychopathology, promote more targeted probing of symptoms, and facilitate differential diagnoses.

Each of the instruments except the K-SADS-P and K-SADS-IVR rate current and past psychopathology. In the K-SADS-PL and K-SADS-E interviews, the presence of all symptoms is queried for both time frames. In the K-SADS-L, however, the majority of affective symptoms are only rated for current episode and past 2 weeks, and all other symptoms are only given lifetime ratings. To date, there have been no formal compari-

<table>
<thead>
<tr>
<th>TABLE 1</th>
<th>Characteristics of the Different K-SADS Interviews</th>
</tr>
</thead>
<tbody>
<tr>
<td>Format</td>
<td>Semistructured</td>
</tr>
<tr>
<td>Integrated parent-child interview</td>
<td>Yes</td>
</tr>
<tr>
<td>Detailed probes</td>
<td>Yes</td>
</tr>
<tr>
<td>Detailed scoring criteria</td>
<td>Yes</td>
</tr>
<tr>
<td>Scales for scoring items</td>
<td>0–3, 0–2</td>
</tr>
<tr>
<td>Impairment ratings</td>
<td>Diagnosis-specific &amp; global</td>
</tr>
<tr>
<td>Skip-out criteria</td>
<td>Yes</td>
</tr>
<tr>
<td>Diagnostic overview</td>
<td>Yes</td>
</tr>
<tr>
<td>Time frame</td>
<td>Present</td>
</tr>
<tr>
<td>Lifetime</td>
<td>Last week</td>
</tr>
</tbody>
</table>

Note: K-SADS = Schedule for Affective Disorders and Schizophrenia for School-Age Children.
*Present and Lifetime version.
*Present Episode version (Chambers et al., 1985).
' Epidemiologic version (H. Orvaschel, 1995, Nova University, unpublished).
sons of the various K-SADS instruments, and psychometric properties are not published for the revised K-SADS-E, K-SADS-L, or K-SADS-IVR interviews. As detailed further in the “Method” and “Results” sections, the K-SADS-PL has many features that make it promising for use in future investigations.

METHOD

Sample. The sample for the K-SADS-PL pilot study consisted of 66 children and adolescents: 55 psychiatric outpatients and 11 normal controls. The mean age of the sample was 12.4 years (SD = 2.6; range = 7 through 17), with the sample approximately evenly divided between patient and control cohorts in age, race, or gender distribution. Consistent with clinical and epidemiological studies (e.g., Costello et al., 1988), however, patients were significantly less likely to be living with both biological parents (20% versus 91%, Fisher’s Exact Test, p < .0001).

Recruitment. Psychiatric patients were recruited from the outpatient clinics at Western Psychiatric Institute and Clinic. Children with below average intelligence or active psychosis were excluded from the study. The normal control children included in this pilot were participants in a larger study on childhood depression (see Ryan et al., 1992, for study description). Informed consent to participate in the K-SADS-PL project was obtained in accordance with the University of Pittsburgh Institutional Review Board guidelines.

Interviewers. The pool of interviewers for this project consisted of four bachelors-level and one masters-level clinician. The four bachelors-level interviewers had worked as a team on the childhood depression study for several years. The one masters-level clinician who participated in data collection was a volunteer. As a semi-structured interview, the K-SADS-PL requires intensive training in the instrument, diagnostic classification, and critical differential diagnostic issues.

Measures

Description of K-SADS-PL. The K-SADS-PL is capable of generating 32 DSM-III-R and DSM-IV Axis I child psychiatric diagnoses. Diagnoses are scored as definite, probable (≥75% of symptom criteria met), or not present. The different components of the K-SADS-PL are described below.

Introductory Interview: The Introductory Interview is used to establish rapport (Rutter and Graham, 1968). It is essential and takes approximately 10 to 15 minutes to complete. In this section, demographic, health, presenting complaint, and prior psychiatric treatment data are obtained, together with information about the child’s school functioning, hobbies, and peer and family relations. Discussion of these latter topics is extremely important, as they provide a context for eliciting mood symptoms (e.g., depression, irritability, anxiety) and obtaining information to evaluate functional impairment.

Screen Interview: One of the main aims of the Screen Interview is to streamline the assessment and enhance administration efficiency. The 82-symptom Screen Interview is divided into 20 different diagnostic areas. (The number of diagnoses assessed with the K-SADS-PL exceeds the number of screen areas, as some diagnostic areas screen for multiple disorders.)

At the conclusion of each diagnostic area, skip-out criteria are delineated for current and past episodes of disorder. The interviewer can skip out of the supplement for a given diagnostic area if the child does not receive a threshold score on any of the symptoms surveyed in that section of the Screen Interview. The diagnostic supplement for a given area is administered if the child receives even one threshold rating. If all skip-out criteria are met, the K-SADS-PL interview is complete after administration of the 82-symptom Screen Interview.

Diagnostic Supplements. The K-SADS-PL has five diagnostic supplements: (1) Affective Disorders; (2) Psychotic Disorders; (3) Anxiety Disorders; (4) Behavioral Disorders; and (5) Substance Abuse, Eating, and Tic Disorders. The skip-out criteria in the Screen Interview specify which section(s) of the supplements, if any, should be completed.

As discussed in the introduction, the Screen Interview is completed before any diagnostic supplements are administered. This provides a diagnostic overview of lifetime psychopathology, allows for more targeted probing of symptoms, and thereby facilitates differential diagnoses. If a child meets possible criteria for two disorders with onset of one preceding the other, the supplement for the diagnosis with the earlier onset is completed first. For example, if a child has evidence of ADHD beginning at age 5, and possible major depressive disorder (MDD) beginning at age 9, the supplement for ADHD would be completed before the supplement for MDD. If the child has a history of attention difficulties associated with ADHD, when inquiring about concentration difficulties in assessing MDD, one would probe to find out whether the onset of depressive symptoms was associated with a worsening of the long-standing concentration difficulties. If there was no change in attention problems with the onset of the depressive symptoms, the symptom concentration difficulties would not be rated positively in the MDD supplement. When the time course of two disorders overlap, supplements for disorders that may influence the course of other disorders are completed first. For example, if there is evidence of substance use and possible mania, the substance abuse supplement would be completed first and care taken to assess the relationship between substance use and manic symptoms.

Time Frame Coding Guidelines. In coding current episodes of disorders, symptoms are rated for the period of maximum severity within the episode. The interviewer is to note in the margins if and when particular symptoms improved or resolved. This permits the interviewer to determine whether the child ever met full diagnostic criteria for the disorder, whether they still meet full criteria, or whether the disorder is in partial remission. For disorders treated with medication (e.g., ADHD), the current ratings describe the most intense severity of symptoms experienced prior to initiation of medication or during "drug holidays." Notes in the margins are used to indicate symptoms targeted effectively with medication. Past diagnoses that are rated in the K-SADS-PL should represent the most severe previous episode. For children with a history of episodic or recurrent disorders, it is recommended that a time line be generated to chart lifetime course of disorder and facilitate scoring of symptoms associated with each episode of illness.

Scoring. The majority of K-SADS-PL items are scored using a 0-3-point rating scale. Scores of 0 indicate no information is available, scores of 1 suggest the symptom is not present, scores of 2 indicate subthreshold levels of symptomatology, and scores of 3...
represent threshold criteria. The remaining items are rated on a 0- to 2-point rating scale.

**Administration Time.** When the K-SADS-PL is administered to normal controls, the parent and child interviews each take approximately 35 to 45 minutes. When the K-SADS-PL is administered to psychiatric patients, depending on the range and severity of psychopathology, parent and child interviews each take approximately 1.25 hours.

**Other Assessments.** In addition to the K-SADS-PL, several rating scales of psychopathology were administered. Measures collected include the Child Behavior Checklist (CBCL) (Achenbach and Edelbrock, 1983), a 113-item parent-report questionnaire which surveys internalizing (e.g., depression, anxiety) and externalizing (e.g., aggression, hyperactivity) symptomatology; the Beck Depression Inventory (BDI) (Beck et al., 1961), a 21-item self-report rating scale of depression for adolescents; the Children's Depression Inventory (CDI) (Kovacs, 1985), a 27-item self-report rating scale of depression for children; the Screen for Child Anxiety Related Emotional Disorders (SCARED) (Birmaher et al., 1997), a new 8-item rating scale with parallel parent (SCARED-P) and child (SCARED-C) report forms which assesses symptoms associated with each of the major child anxiety disorder diagnoses; and the Conners Abbreviated Questionnaire/Parent version (Conners and Barkley, 1985), a 10-item scale which assesses ADHD symptoms.

**Procedure.** With children, the K-SADS-PL was administered by interviewing parent(s) first, then interviewing the child alone. With adolescents, the order of administration was reversed. Interviews with both informants were completed by the same interviewer, and summary DSM-IV diagnoses were assigned by using composite ratings which synthesized the parent and child data. Two coders checked to verify accurate utilization of DSM-IV criteria in the assignment of final diagnoses. Interrater and test-retest reliability interviews were checked by independent coders blind to the summary diagnoses verified by the other coder.

**Interrater Reliability.** Interviews of 15 randomly selected subjects were audiotaped for purposes of obtaining interrater reliability estimates. In rating audiotapes, interviewers were blind to the results of the initial interview and all other information about the child. Of the subjects included in the interrater reliability data pool, 10 were patients and 5 were controls. The mean age of these subjects was 12.1 (SD = 2.8; range = 8 through 16); 4 (27%) were female, and 13 (87%) were Caucasian.

**Test-Retest Reliability.** Twenty subjects were randomly selected for reinterview for purposes of obtaining test-retest reliability data. Retest interviews were conducted a mean of 17.9 days (range = 2 to 38 days) after initial interviews. Fifty percent of the interviews were completed within 2 weeks, 75% within 1 month. Test-retest interviews were also completed blind to results of the initial interview and all other information about the child. Nineteen patients and one control subject were included in the test-retest reliability pool. The mean age of these subjects was 12.2 (SD = 3.1; range = 7 through 17); 14 (70%) were female, and 13 (65%) were Caucasian.

**Concurrent Validity.** To determine concurrent validity (n = 66) of the skip-out criteria and diagnoses generated with the K-SADS-PL, children were divided into groups based on whether they screened positively for or met criteria for (1) any depressive disorder, (2) ADHD, (3) any anxiety disorder, and (4) any behavioral disorder. Assignment to each of the groups was not mutually exclusive, as the majority of children met criteria for multiple diagnostic categories. The CDI, BDI, and CBCL Internalizing scores were used as indices of concurrent validity for depression; the Conners rating scale was used as an index of concurrent validity for ADHD; the SCARED-P, SCARED-C, and CBCL Internalizing scale were used as indices of concurrent validity for anxiety; and the CBCL Externalizing scale was used as an index of concurrent validity for any behavioral disorder.

**Statistical Analyses**

Complete data were available for most subjects. Two (3%) parents failed to complete the parent reports (CBCL, SCARED-P, Conners), and one (1.5%) adolescent failed to complete the BDI. When examining psychometric properties of diagnoses assigned with the K-SADS, definite and probable (≥75% of criteria for diagnosis met) diagnostic categories were collapsed. Interrater (n = 15) and test-retest (n = 20) reliability data are reported for (1) skip-out criteria in the Screen Interview and (2) diagnoses derived by synthesizing parent and child K-SADS-PL data. Percent agreement was used to generate interrater reliability estimates, as there were an insufficient number of cases (n < 5) to justify calculation of a k statistic (Cohen, 1960) in most diagnostic categories. In the test-retest reliability data set, several diagnoses were screened and assigned with adequate frequency (e.g., at least five cases at time 1 or time 2) to justify calculation of the k statistic. Criteria proposed by Landis and Koch (1977) were used to interpret the k coefficients: excellent reliability, k > .75; good reliability, k = .59 to .75; fair reliability, k = .40 to .58; and poor reliability, k < .40.

Prior to conducting analyses of the concurrent validity data, the CDI and BDI raw scores were subjected to z score transformation to allow for examination of depression scores for children and adolescents simultaneously. The normality of the distribution of scores on all measures was also examined. As all the data were normally distributed, t tests were used to examine group differences on rating scale data. Levene F test was used to determine homogeneity of variance, and t tests for unequal variances are reported when appropriate.

**RESULTS**

**Diagnostic Profile of the Sample.** The majority of patients (n = 55) met criteria for multiple current (2.8 ± 1.4, range = 1 to 6) and past (1.1 ± 1.5, range = 0 to 8) diagnoses. The most frequently occurring present diagnoses assigned were ADHD (42%), oppositional defiant disorder (34%), MDD (23%), bipolar disorders (19%), PTSD (17%), generalized anxiety disorder (17%), dysthymia (11%), and simple phobia (11%). Only a small proportion of children had "pure" diagnostic profiles. Four (6%), 3 (5%), and 12 (18%) of the children met criteria for pure affective, pure anxiety, or pure behavioral disorders, respectively.

**Skip-Out Criteria**

Of the 20 diagnostic areas surveyed in the Screen Interview, each of the patients screened positive for an average of 3.7 (median/mode = 3; range = 1 to 8) current and 2.6 (median = 2; mode = 1; range = 0 to 10) past possible diagnoses. Eight (73%) of the 11 normal
subjects skipped out of all diagnostic areas, 1 screened positive for a possible current disorder, and 3 screened positive for a possible past disorder.

**Interrater Reliability** \((n = 15)\). Interrater reliability was examined across the 20 diagnostic areas surveyed in the Screen Interview. The average agreement in the utilization of skip-out criteria across the 20 diagnostic areas was 99.7% \((\text{range} = 93\% \text{ to } 100\%)\) for the assessment of current diagnoses and 100% for past diagnoses.

**Test-Retest Reliability** \((n = 20)\). Of the current diagnostic areas were screened and later diagnosed with sufficient frequency to calculate the \(\kappa\) statistic: depressive disorders \((n = 16; \kappa = .52)\); generalized anxiety disorder \((n = 11; \kappa = .70)\); PTSD \((n = 15; \kappa = .56)\); ADHD \((n = 10; \kappa = .59)\); and oppositional defiant disorder \((n = 8; \kappa = .50)\). These \(\kappa\) values are in the fair range.

**Concurrent Validity** \((n = 66)\). Children who screened positive for current depression \((n = 37)\) scored significantly higher than the other children \((n = 29)\) on the \(z\) score-transformed depression \((\text{DEP/POS} = 0.21 \pm 1.0)\); \(\text{DEP/NEG} = -0.33 \pm 0.9; t_{63} = 2.29, p < .03\) and CBCL Internalizing \((\text{DEP/POS} = 67.5 \pm 9.7); \text{DEP/NEG} = 55.6 \pm 14.4; t_{45.4} = 3.75, p < .0005)\) scales. Children who screened positive for current ADHD \((n = 36)\) also scored higher than the other children \((n = 30)\) on the Conners Parent Rating Scale \((\text{ADHD/POS} = 22.6 \pm 14.3); \text{ADHD/NEG} = 13.7 \pm 16.7; t_{61} = 2.26, p < .03\). In addition, children who screened positive for any current anxiety disorder \((n = 49)\) scored significantly higher than the other children \((n = 17)\) on the SCARED-P \((\text{ANX/POS} = 22.1 \pm 14.5); \text{ANX/NEG} = 7.6 \pm 9.0; t_{46.4} = 4.76, p < .0001), \text{SCARED-C} \((\text{ANX/POS} = 22.7 \pm 14.2); \text{ANX/NEG} = 10.5 \pm 8.4; t_{64} = 3.33, p < .0001\), and CBCL Externalizing \((\text{ANX/POS} = 65.2 \pm 11.5); \text{ANX/NEG} = 54.4 \pm 14.9; t_{61} = 3.05, p < .003)\) scales; and the children who screened positive for any current behavioral disorder \((n = 45)\) scored significantly higher than the other children \((n = 21)\) on the CBCL Externalizing scale \((\text{BEH/POS} = 61.1 \pm 9.9); \text{BEH/NEG} = 51.7 \pm 9.2; t_{61} = 5.90, p < .0001)\).

**Diagnoses**

**Interrater Reliability** \((n = 15)\). In the interrater reliability data pool, 10 current and 14 lifetime diagnoses were assigned, with each diagnosis on average being assigned to two to three children each. Percent agreement in assigning present and lifetime diagnoses were both 98% \((\text{range} = 93\% \text{ to } 100\%)\).

**Test-Retest Reliability** \((n = 20)\). Table 2 depicts the test-retest reliability estimates for the diagnoses as-

<p>| Test-Retest Reliability of K-SADS-PL Diagnoses ((n = 20)) |
|-----------------------------|-----------------------------|</p>
<table>
<thead>
<tr>
<th>Present Diagnoses</th>
<th>Lifetime Diagnoses</th>
</tr>
</thead>
<tbody>
<tr>
<td>n</td>
<td>(\kappa)</td>
</tr>
<tr>
<td><strong>Major depressive disorder</strong></td>
<td>10</td>
</tr>
<tr>
<td>Any depressiona</td>
<td>11</td>
</tr>
<tr>
<td>Depressive disorder NOS</td>
<td>2</td>
</tr>
<tr>
<td>Any bipolar disorderb</td>
<td>5</td>
</tr>
<tr>
<td>Generalized anxiety disorder</td>
<td>8</td>
</tr>
<tr>
<td>Posttraumatic stress disorder</td>
<td>8</td>
</tr>
<tr>
<td>Any anxiety disorderc</td>
<td>11</td>
</tr>
<tr>
<td>ADHD</td>
<td>7</td>
</tr>
<tr>
<td>Conduct disorder</td>
<td>2</td>
</tr>
<tr>
<td>Oppositional defiant disorder</td>
<td>6</td>
</tr>
</tbody>
</table>

**Note:** K-SADS-PL = Schedule for Affective Disorders and Schizophrenia for School-Age Children-Present and Lifetime version; NOS = not otherwise specified; ADHD = attention-deficit hyperactivity disorder.

* *Any depression* is defined as major depressive disorder and/or dysthymia.

* *Any bipolar disorder* is defined as bipolar I or bipolar NOS.

* *Any anxiety disorder* is defined as panic, separation anxiety, social phobia, agoraphobia, simple phobia, generalized anxiety, or obsessive-compulsive disorder.

* Three of the four discrepant cases of ADHD were between assigning a probable diagnosis and no diagnosis. If the algorithm for considering a diagnosis positive was altered, such that only definite cases were counted, the test-retest reliability \(\kappa\) for lifetime diagnosis of ADHD \((n = 5)\) would have been .86.
<table>
<thead>
<tr>
<th>Retest interval</th>
<th>K-SADS-PL</th>
<th>K-SADS-P^</th>
<th>CAS^</th>
<th>ISC^</th>
<th>DICA-C</th>
<th>DISC-2.1</th>
<th>CAPA-C</th>
</tr>
</thead>
<tbody>
<tr>
<td>Major depressive disorder</td>
<td>1.00</td>
<td>1.00</td>
<td>.90</td>
<td>.78</td>
<td>.63</td>
<td>.74</td>
<td>—</td>
</tr>
<tr>
<td>Any depression</td>
<td>—</td>
<td>—</td>
<td>.90</td>
<td>.78</td>
<td>.77</td>
<td>.70</td>
<td>—</td>
</tr>
<tr>
<td>Generalized/overanxious disorder</td>
<td>.83</td>
<td>.38</td>
<td>.81</td>
<td>.72</td>
<td>.43</td>
<td>.66</td>
<td>—</td>
</tr>
<tr>
<td>Any anxiety disorder</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>ADHD</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Oppositional defiant disorder</td>
<td>.90</td>
<td>.83</td>
<td>.90</td>
<td>.79</td>
<td>.64</td>
<td>.50</td>
<td>.61</td>
</tr>
</tbody>
</table>

Note: K-SADS-PL = Schedule for Affective Disorders and Schizophrenia for School-Age Children-Present and Lifetime version; ADHD = attention-deficit hyperactivity disorder.

^K-SADS-Present Episode (Chambers et al., 1985).
^aChild Assessment Schedule (Hodges et al., 1982).
^bInterview Schedule for Children (Kovacs, 1985).
^cDiagnostic Interview for Children and Adolescents (Welner et al., 1987).
^dDiagnostic Interview for Children Version 2.1 (Jensen et al., 1995).
^eChild and Adolescent Psychiatric Assessment (Angold and Costello, 1995).
^f"Any depression" is defined as major depressive disorder and/or dysthymia.
^g"Any anxiety disorder" is defined as panic, separation, anxiety, social phobia, agoraphobia, simple phobia, generalized anxiety, or obsessive-compulsive disorder.

Test-retest reliability \( \kappa \) coefficients were in the excellent range for most present and lifetime diagnoses. Test-retest reliability \( \kappa \) coefficients were in the excellent range for lifetime diagnoses of MDD, any depression, depressive disorder not otherwise specified (NOS), any bipolar disorder, generalized anxiety disorder, conduct disorder, and oppositional defiant disorder, and in the good range for lifetime diagnoses of PTSD and any anxiety disorder.

**Concurrent Validity** (n = 66). Children who met criteria for current depressive disorders (n = 27) scored significantly higher than the other children (n = 39) on the \( z \) score-transformed depression (DEP/POS = 0.32 ± 1.1; DEP/NEG = −0.26 ± 0.8; \( t_{60} = 2.47, p < .01 \)) and CBCL Internalizing (DEP/POS = 68.7 ± 10.0; DEP/NEG = 58.4 ± 13.6; \( t_{62.4} = 3.54, p < .001 \)) scales. Children who who met criteria for current ADHD (n = 30) also scored higher than the other children (n = 36) on the Conners Parent Rating Scale (ADHD/POS = 25.8 ± 13.8; ADHD/NEG = 13.1 ± 15.7; \( t_{63} = 3.43, p < .001 \)). In addition, children who who met criteria for any current anxiety disorder (n = 34) scored significantly higher than the other children (n = 32) on the SCARED-P (ANX/POS = 25.5 ± 14.6; ANX/NEG = 11.4 ± 11.4; \( t_{60} = 4.32, p < .0001 \)), SCARED-C (ANX/POS = 26.1 ± 14.7; ANX/NEG = 12.6 ± 9.3; \( t_{56.4} = 4.47, p < .0001 \)), and CBCL Internalizing (ANX/POS = 66.8 ± 11.5; ANX/NEG = 58.1 ± 13.7; \( t_{60} = 2.76, p < .01 \)) scales; and children who met criteria for any current behavioral disorder (n = 38) scored significantly higher than the other children (n = 28) on the CBCL Externalizing scale (BEH/POS = 68.7 ± 8.6; BEH/NEG = 53.6 ± 10.7; \( t_{63} = 6.29, p < .0001 \)).

**Comparison With Other Diagnostic Instruments.** Table 3 compares the test-retest reliability of the K-SADS-PL and several other child diagnostic interviews. Test-retest reliability estimates in all studies were derived from samples of psychiatric patients approximately the same age as the children in this study. Diagnoses generated with the K-SADS-P, Child Assessment Schedule, Interview Schedule for Children, and Diagnostic Interview Schedule for Children Version 2.1 were derived by synthesizing parent and child reports, and diagnoses assigned with the Diagnostic Interview for Children and Adolescents and the Child and Adolescent Psychiatric Assessment were derived from children's reports alone. Although the test-retest interval in this study is significantly greater than the interval in the other reports, the K-SADS-PL compares favorably with the other instruments and appears to have some advantage over several of the other instruments in the assessment of affective and anxiety disorders.
DISCUSSION

The K-SADS-PL (1) assesses current and lifetime psychiatric history; (2) surveys additional disorders not assessed in previous versions of the K-SADS (e.g., ADHD, PTSD, tic disorders); (3) contains improved probes and anchor points for scoring previously problematic items; (4) includes diagnosis-specific impairment ratings to assist in "caseness" determination; (5) generates DSM-III-R and DSM-IV diagnoses; and (6) divides the symptoms surveyed in the instrument into an 82-symptom Screen Interview and five supplements to facilitate differential diagnoses and expedite administration with patients and normal controls.

The K-SADS-PL is an integrated parent-child interview. Diagnoses are generated by synthesizing parent and child data, with greater weight typically given to parents' reports of observable behavior and children's reports of subjective experiences (Herjanic and Reich, 1982). While there is consensus in the field concerning the importance of obtaining information from multiple informants and including children in the data collection process when assessing child psychopathology (Achenbach et al., 1987; Jensen et al., 1995; Reich and Earls, 1987; Rutter and Graham, 1968), more research is needed to determine the optimal way to combine data from multiple informants (Offord, 1995; Young et al., 1987).

Results of this initial test of the psychometric properties of the K-SADS-PL are quite promising. The interrater reliability data collected in this study were excellent and comparable with those reported by other investigators using both semistructured and fully structured child diagnostic interviews (Herjanic and Reich, 1982; Hodges et al., 1982; Shaffer et al., 1993). Agreement on the scoring of the skip-out criteria determining the need to complete the diagnostic supplements approached 100%. The raters disagreed on the need to complete a supplement in only one case, and in that case, no diagnosis was assigned by the rater who went on to complete the items in the diagnostic supplement. Interrater agreement was also very high in the assignment of diagnoses. Interviewers agreed 100% of the time in the assignment of MDD, dysthymia, depressive disorder NOS, bipolar I disorder, social phobia, generalized anxiety disorder, and PTSD. Agreement on all other assigned diagnoses was 93%.

Test-retest reliability on the scoring of the skip-out criteria was quite high. Test-retest reliability estimates for the assignment of diagnoses were in the excellent to good range for most present and lifetime diagnoses. Test-retest reliability \( \kappa \) coefficients were in the excellent range for present and/or lifetime diagnoses of MDD, any depression, depressive disorder NOS, any bipolar disorder, generalized anxiety, any anxiety, conduct, and oppositional defiant disorder (.77 to 1.00) and in the good range for present diagnoses of PTSD and ADHD (.63 to .67). As depicted in Table 3, the K-SADS-PL does notably better than its predecessor, the K-SADS-P, in reliably diagnosing affective and anxiety disorders. In addition, although the test-retest interval was greater in this study than in other published reports, the K-SADS-PL fares quite favorably when compared with other child psychiatric interviews and has some advantages over several of the other instruments in the assessment of affective and anxiety disorders.

The concurrent validity of both the skip-out criteria and the diagnoses generated with the K-SADS-PL was well supported. In all cases, children who screened positive or met criteria for a specific diagnostic category (e.g., depressive disorder, ADHD, anxiety disorder, behavioral disorder) scored significantly higher than the other children in the study on the rating scales assessing the symptoms associated with that particular diagnostic category.

While the rating scale data provide preliminary support for the validity of the diagnoses generated with the K-SADS-PL, determination of diagnostic validity is a very complicated endeavor, as there is no "gold standard" against which to compare the K-SADS-PL diagnoses (Achenbach et al., 1987; Hodges, 1994). Rating scales are relatively insensitive in classifying children with particular disorders (Bird et al., 1991; Hodges...
et al., 1982), and there are limitations to diagnoses derived from charts (Welner et al., 1987) and from clinical interviews (Robins, 1985; Steiner et al., 1995) as well. The collection of family history or longitudinal follow-up data would have been preferable to validate the diagnoses (Feighner et al., 1972; Malgady et al., 1992), but was clearly beyond the scope of the present investigation.

There are a number of other limitations to the present study, in addition to those of the validity data. The test-retest reliability sample was small, prohibiting the examination of age (Edelbrock et al., 1985; Schwab-Stone et al., 1994) and race (Segal et al., 1994) effects on the psychometric properties of the K-SADS-PL. In addition, the children in the study met criteria for only a limited range of disorders, precluding study of the ability of the K-SADS-PL to diagnose less common child psychiatric conditions. The findings of this study can also not be generalized to population samples, as the data in the literature clearly show that reliability estimates tend to be higher in clinical versus community samples (Boyle et al., 1993; Jensen et al., 1995).

Despite these limitations, the K-SADS-PL appears quite promising for use in future investigations. The results of this study suggest the K-SADS-PL generates reliable and valid psychiatric diagnoses, with particular strengths in assessing affective and anxiety disorders. Although the concurrent validity data presented in this report support the independent use of the K-SADS-PL for purposes of generating psychiatric diagnoses in children, we recommend the K-SADS-PL be utilized as part of a comprehensive assessment battery which includes rating scale data from parents, children, and teachers, whenever possible. Consistent with the views of others, we believe the most valid diagnoses integrate data from all available sources by using either the "best estimate" (Leckman et al., 1982) or PLASTIC (prospective, longitudinal, all source, treatment, impairment, and clinical presentation) (Young et al., 1987) methods.

REFERENCES

Robins L (1985), Epidemiology: reflection of testing the validity of psychiatric interviews. Arch Gen Psychiatry 42:918–924


